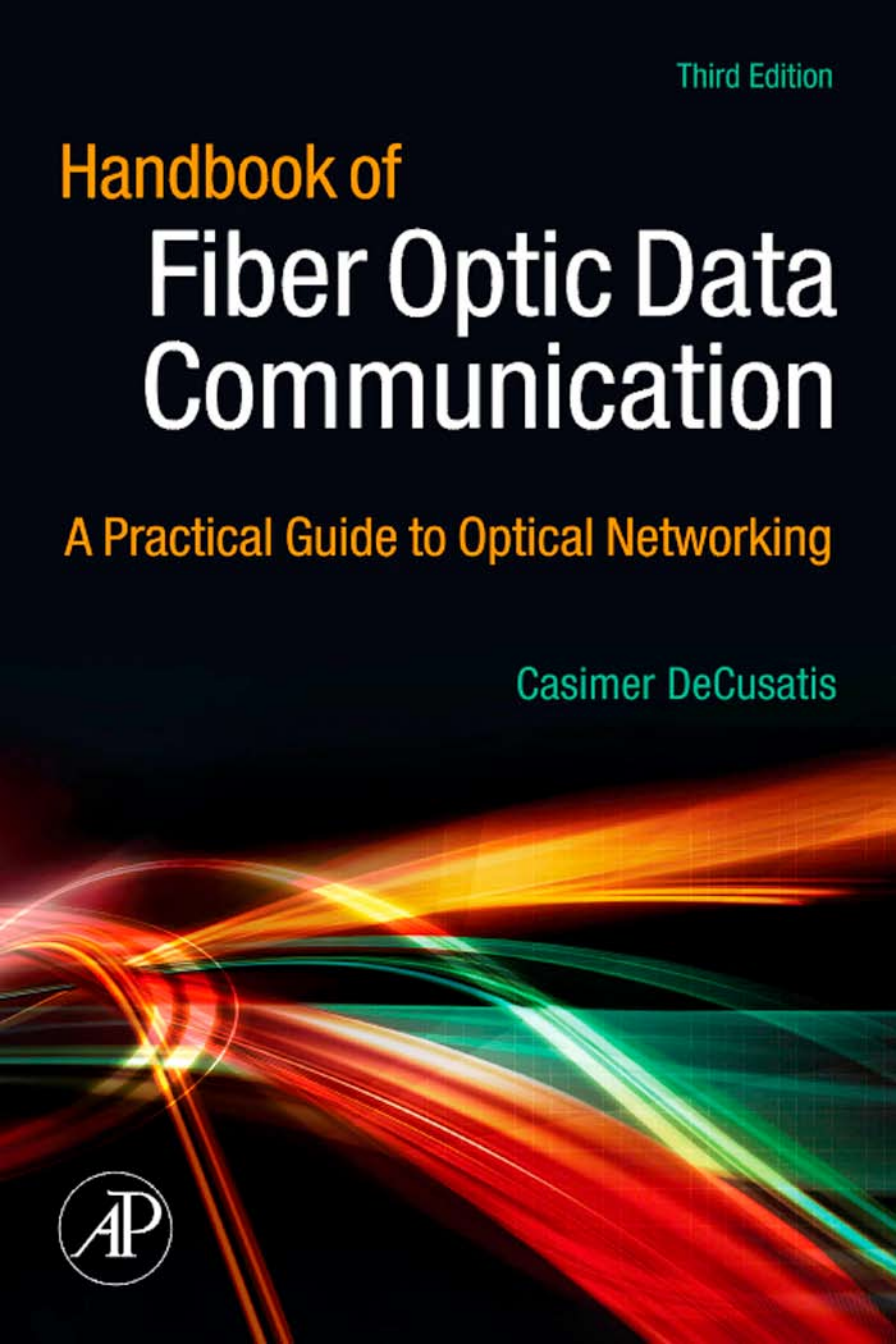
Third Edition

Handbook of Fiber Optic Data Communication

A Practical Guide to Optical Networking

Casimer DeCusatis



Handbook of Fiber Optic Data Communication

This page intentionally left blank

Handbook of Fiber Optic Data Communication

A Practical Guide to Optical Networking

Third Edition

Edited by Casimer DeCusatis

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Elsevier Academic Press
Design Direction: Joanne Blank
Cover Design: Gary Ragaglia
Cover Images© iStockphoto
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper.

Copyright © 2008, 2002, 1998, Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone: (+44) 1865 843830, fax: (+44) 1865 853333, E-mail: permissions@elsevier.co.uk. You may also complete your request on-line via the Elsevier homepage (http://elsevier.com), by selecting "Customer Support" and then "Obtaining Permissions."

Library of Congress Cataloging-in-Publication Data
Handbook of fiber optic data communication : a practical guide to optical networking / editor, Casimer DeCusatis.—3rd ed.
p. cm.
Includes index.
ISBN 978-0-12-374216-2 (hbk. : alk. paper) 1. Optical communications. 2. Fiber optics. 3. Data transmission systems. I. DeCusatis, Casimer.
TK5103.59.H3515 2008
621.39′81—dc22
2007045797

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-374216-2

For all information on all Elsevier Academic Press publications visit our Web site at www.books.elsevier.com

Printed in The United States of America
08 09 10 11 12 9 8 7 6 5 4 3 2 1

Working together to grow libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Table of Contents

Part I Technology Building Blocks 1

Chapter 1 Computers Full of Light: A Short History of Optical Data Communication (J. Hecht) 3

Chapter 2 Optical Fiber, Cables, and Connectors (U. Osterberg) 19

Chapter 3 Small Form Factor Fiber-Optic Interfaces (J. Fox, C. DeCusatis) 43

Chapter 4 Specialty Fiber-Optic Cables (J. Fox, C. DeCusatis) 63

Case Study: Multimode Fiber Reuse for High-Speed Storage Area Networks 87

Chapter 5 Optical Sources: Light-Emitting Diodes and Laser Technology (W. Jiang, M. S. Lebby) 91

Chapter 6 Detectors for Fiber Optics (C. J. S. DeCusatis, C. Jiang) 133

Chapter 7 Receiver Logic and Drive Circuitry (R. Sundstrom, E. Maass, contributions by S. Kipp) 163

Chapter 8 Optical Subassemblies (H. Stange) 177

Chapter 9 Alignment Metrology and Manufacturing (D. P. Clement, R. C. Lasky, and D. Baldwin) 193

Chapter 10 Packaging Assembly Techniques (G. Raskin) 219

Part II Links and Network Design 239

Chapter 11 Fiber-Optic Transceivers (M. Langemwalter and R. Johnson, contributions by R. Atkins) 241

Chapter 12 Optical Link Budgets and Design Rules (C. DeCusatis) 271

Case Study: WDM Link Budget Design 305

Chapter 13 Planning and Building the Optical Link (R. T. Hudson, D. R. King, T. R. Rhyne, T. A. Torchia) 307

Chapter 14 Test and Measurement of Fiber Optic Transceivers (G. LeCheminant) 339

Chapter 15 Optical Wavelength Division Multiplexing for Data Communication Networks (K. Grobe) 371

Case Study: National LambdaRail Project 399

Case Study: Optical Networks for Grid Computing 403

Chapter 16 Passive Optical Networks (K. Grobe) 405

Part III Applications & Industry Standards 425

Chapter 17 Optical Interconnects for Clustered Computing Architectures (D. B. Sher, C. DeCusatis) 427

Case Study: Parallel Optics for Supercomputer Clustering 453

Chapter 18 Manufacturing Environmental Laws, Directives, and Challenges (J. Quick) 455

Chapter 19 ATM, SONET, and GFP (C. Beckman, R. Thapar) 473

Case Study: Facilities-Based Carrier Network Convergence and Bandwidth on Demand 503

Chapter 20 Fibre Channel—The Storage Interconnect (S. Kipp, A. Benner) 505

Case Study: Design of Next Generation I/O for Mainframes 533

Case Study: Storage Area Network (SAN) Extension for Disaster Recovery 535

Chapter 21 Enterprise System Connection (ESCON) Fiber-Optic Link (D. J. Stigliani) 537

Case Study: Calculating an ESCON Optical Link Budget 567

Chapter 22 Enhanced Ethernet for the Data Center (C. DeCusatis) 571

Case Study: Unbundling the Local Loop for Triple Play Networks 591

Chapter 23 FDDI and Local Area Networks (R. Thapar) 593

Chapter 24 InfiniBand—The Cluster Interconnect (A. Ghiasi) 605

Part IV Emerging Technologies & Industry Directions 629

Chapter 25 Emerging Technology for Fiber-Optic Data Communication (C. S. Li) 631

Case Study: Customer-Owned Wavelengths and P2P Optical Networking 653

Chapter 26 Optical Backplanes, Board and Chip Interconnects (R. Michalzik) 657

Chapter 27 Silicon Photonics (N. Izhaky) 677

Chapter 28 Nanophotonics and Nanofibers (L. Tong, E. Mazur) 713

Appendix A Measurement Conversion Tables 729

Appendix B Physical Constants 731

Appendix C The 7-Layer OSI Model 733

Appendix D Network Standards and Data Rates 735

Appendix E Other Datacom Developments 751

Appendix F Laser Safety 755

Index 763

This page intentionally left blank

Preface to the Third Edition

SONET[1] on the Lambdas[2]

(by C. DeCusatis, with sincere apologies to Milton[3])

When I consider how the light is bent
By fibers glassy in this Web World Wide,
Tera- and Peta-, the bits fly by
Are they from Snell and Maxwell sent
Or through more base physics, which the Maker presents
(lambdas of God?) or might He come to chide
"Doth God require more bandwidth, light denied?"
Consultants may ask; but Engineers to prevent
that murmur, soon reply "The Fortune e-500 do not need
mere light alone, nor its interconnect; who requests
this data, if not clients surfing the Web?" Their state
is processing, a billion MIPS or CPU cycles at giga-speed.
Without fiber-optic links that never rest,
The servers also only stand and wait.

As this book goes to press, I am pleased to say that the world of optical data communication is well established and continues to thrive. Mature technologies combined with high-volume, low-cost manufacturing have made high-performance optical data links more affordable than ever before and have turned some of the early technologies into commodities. Applications for fiber-optic networking have grown significantly. This goes beyond Internet and Web traffic to encompass areas such as disaster recovery, video distribution, massively parallel clustered computing, and networked storage. (Large corporations now boast multi-terabyte, petabyte, or even exabyte databases interconnected with their core business functions.) The distinction between datacom and telecom technologies continues to blur, with the encapsulation of traditional data center protocols over

[1]Synchronous Optical Network.
[2]The Greek symbol "lambda" or λ is commonly used in reference to an optical wavelength.
[3]The original author of the classic sonnet "On His Blindness."

metropolitan and wide area networks designed for voice traffic. Network convergence and the triple or quadruple play for service providers have entered common usage, but the unique requirements of data communication networks remain (including very low error rates, long unrepeated distances, ease of use for untrained staff, and an unprecedented combination of high reliability and low cost in demanding environments). These many developments, coupled with the continued success of previous editions, led to the decision that the time was right to update this Handbook once again.

Since the first edition was published over 10 years ago, I have tried to continually incorporate feedback and comments from readers to improve this book and ensure that it continues to provide a single, indispensable reference for the optical data communication field. Previous editions had experimented with a two-volume set of Handbooks. But you, the readers who make use of this book every day, have consistently emphasized the importance of having a single volume as your one-stop reference source. In this edition, I have taken your advice and have returned the Handbook to its original design. This one book contains an overview of the entire optical data communication field, broken down into basic technology, link design, planning, installation, testing, protocols, applications, and future directions.

A great deal of new material has been added, and many familiar chapters have been updated to reflect new types of optical components, connectors, cables, and other devices. Some legacy applications that are not as widely used have been edited to their essential material only, such as FDDI and ESCON. Others have been expanded, and we have added the latest updates to Fibre Channel/FICON, InfiniBand, and SONET/SDH. Some technologies that were just emerging when the previous edition was published are now commonplace; among these are pluggable small form factor transceivers. Completely new chapters deal with issues that did not exist when the last edition was published, including Enhanced Ethernet for the data center, silicon photonics, and nanofibers. Throughout I have tried to maintain a focus on practical applications. This edition includes about a dozen case studies that either provide numerical examples of the principles discussed in the text or discuss real-world applications using grid computing, triple-play networks, optically interconnected supercomputers, and other areas. Our industry is just beginning to see the promise of all-optical networking emerge—application-neutral, distance-independent, infinitely scalable, user-centric networks that catalize real-time global computing, advanced streaming multimedia, distance learning, telemedicine, and a host of other applications. We hope that those who build and use these networks will benefit in some measure from this book.

An undertaking such as this would not be possible without the concerted efforts of many contributing authors and the publisher's supportive staff, to all of whom

I extend my deepest gratitude. As always, this book is dedicated to my mother and father, who first helped me see the wonder in the world; to the memory of my godmother Isabel; and to my wife, Carolyn, and my daughters Anne and Rebecca, without whom this work would not have been possible.

Dr. Casimer DeCusatis, Editor
Poughkeepsie, New York
August 2007

This page intentionally left blank

Part I

Technology Building Blocks

This page intentionally left blank

1

Computers Full of Light: A Short History of Optical Data Communications

Jeff Hecht
Consultant, Auburndale, MA.

To those of us who grew up in the electronics era, optical communications is a new technology. But if you look back, you can find that the age of telecommunications started not with the well-known electrical telegraph, but with optical telegraphs that first came into use in the late eighteenth century. The new age of optical communications has been powered by two new technologies invented in the mid-twentieth century—lasers and fiber optics.

The shift to optics coincided with the change from analog to digital transmission in the telephone network and with the growing importance of computer data transmission. Historians of technology state that technology evolves and that evolution is evident in the changes that have combined optical and digital technology, both on large and small scales in the global telecommunications network.

1.1 THE OPTICAL TELEGRAPH

The idea of telegraphing signals to remote locations emerged long before scientists had any idea how to control electricity. The first telegraph proposals were for semaphore-based systems that relayed signals between a series of stations. The operator of one would spell out a message as a series of characters, which the operator of the next would view through a telescope, write down, and relay to the operator of the next. The scheme was labor-intensive, but at the time labor was cheap, and it could send signals much faster than horses.

The oldest recorded proposal for an optical telegraph dates from March 21, 1684, when English scientist Robert Hooke described "a way how to communicate one's mind at great distances" to fellow members of London's Royal Society. Hooke suggested that the towers display light-colored characters at night and dark

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

ones during the day, so that they could be easily seen, and he proposed coding the symbols to prevent eavesdropping.[1] It was a remarkably prescient idea, but it would take a century before the first practical system was built.

The impetus for success came from the French Revolution, which left France in turmoil and surrounded by enemies. Optical telegraphs had been demonstrated by then, but only over short distances. Claude Chappe and his four brothers set themselves to the far more ambitious task of building a national optical telegraph network. After some false starts, in March 1791 they succeeded in sending signals between two French towns and made a point of having local officials confirm the demonstration.

The Chappe brothers then asked the revolutionary government to fund their plans to build an optical telegraph network. Claude moved his experiments to Paris, and his brother Ignace was elected to the new Legislative Assembly, where he became a member of the Committee for Public Invention. Those connections helped the Chappes gain support as they refined their technology. First they tested a pulley-driven array of five sliding panels that offered 32 possible combinations, enough to spell the alphabet plus a few other symbols. Later they shifted to a semaphore with two arms on the ends of a longer horizontal beam, as shown in Fig. 1.1.

To prove their design would work, the Chappes built a demonstration system spanning two segments, one of 15 kilometers (km) and the second of 11 km, and on July 12, 1793, they transmitted a 26-word message in 11 minutes, incredibly fast by the standards of the time.[2] Two weeks later the government agreed to build a 15-station line spanning 120 km from Paris to Lille. That system began operating less than a year later and grew steadily because the war-torn country needed to keep in touch with its frontiers. The system survived the fall of Napoleon and the restoration of Louis XVIII, and ultimately other countries built their own optical telegraphs, as Gerard J. Holzmann and Björn Pehrson recount in a fascinating book titled *The Early History of Data Networks*.[3]

Optical telegraphs launched the age of telecommunications, but by the 1830s a competitor had emerged—the electrical telegraph. The new electrical systems were cheaper to build and operate and could transmit signals at any time, not just when the sun was shining and the air was clear.

Optical communications was not entirely forgotten in the years that followed. In 1880, Alexander Graham Bell demonstrated the “Photophone,” an optical version of the telephone that modulated the intensity of reflected sunlight with voice signals. The Photophone fascinated Bell, but it could not compete with his earlier

[1]Gerard J. Holzmann and Björn Pehrson, *The early history of data networks* (Los Alamitos, Calif.: IEEE Computer Society Press, 1995), pp. 35–38.

[2]Ibid., p. 61.

[3]Holzmann and Pehrson, *The early history of data networks*.

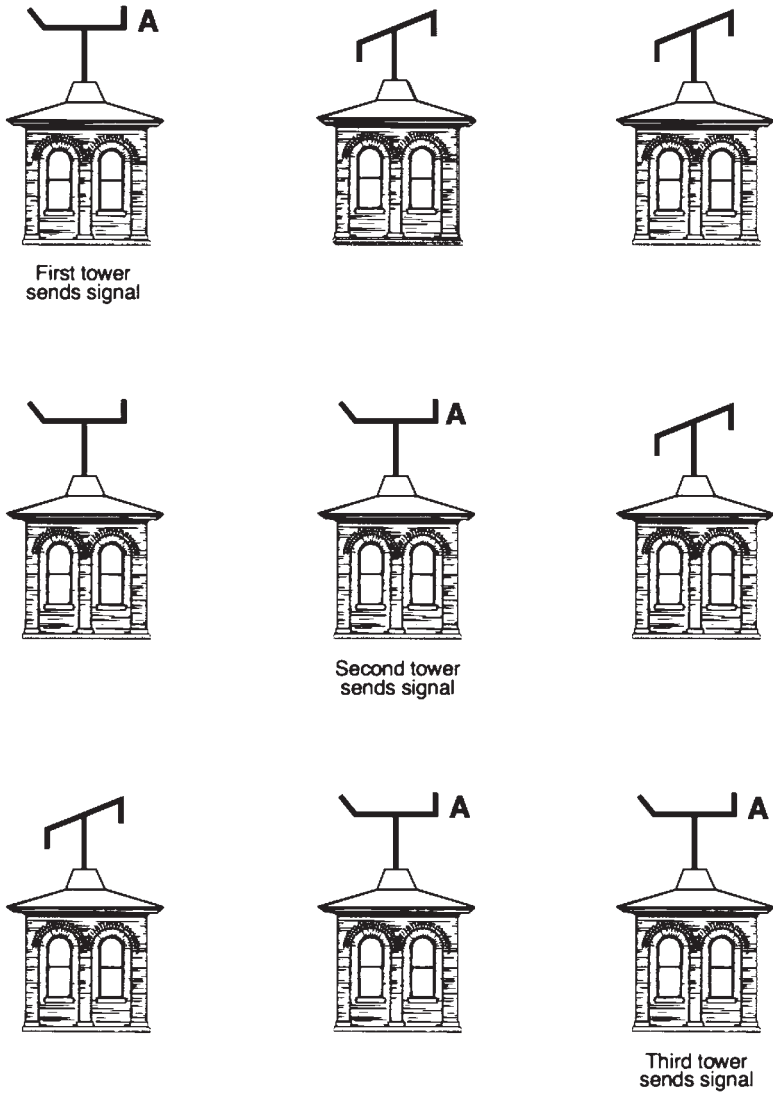


Figure 1.1 Signal transmission along a series of Chappe-style optical telegraph towers.

invention, the wired telephone. Like the electrical telegraph, the wired phone could transmit signals day or night, regardless of the weather.[4]

1.2 LASERS REVIVE OPTICAL COMMUNICATIONS

The birth of the laser launched the new age of optical communications. The first step on the path to the laser was the 1954 invention of its microwave counterpart, the maser, by Charles Townes, then at Columbia University. The amplification of stimulated emission from material contained in a resonant cavity made the maser oscillate at the frequency of the stimulated emission. Importantly, maser output was coherent and limited to a narrow range of frequencies.

The next logical step was to extend the maser principle to the much higher frequencies of light waves. The team of Townes and Arthur Schawlow and, separately, Gordon Gould, working by himself, both proposed similar designs for a laser, essentially solving the same physics problem and coming out with the same answer. However, it was Theodore Maiman, working at Hughes Research Laboratories in California, who succeeded in making the first laser on May 16, 1960.[5]

Optical communications was a key application envisioned by laser developers. As a coherent oscillator, the laser was analogous to the coherent oscillators used in radio communications, but because light waves had much higher frequencies, they promised much higher transmission capacity. Maiman's demonstration opened the floodgates to a series of experiments, first with the ruby laser Maiman had invented and later with the helium-neon gas laser invented at Bell Labs. Initial tests showed that laser beams could be modulated in intensity to carry a signal and that they could travel many miles through clear air. However, further tests eventually revealed that fog, clouds, or precipitation could attenuate or block the beam, making long-distance signal transmission unreliable through open air.

Short laser links through the air did work reasonably well. The National Aeronautics and Space Agency (NASA) considered them to replace umbilical communication cables connecting spacecraft waiting for launch with mission control. Businesses considered lasers for short links through the air between buildings that did not require the Federal Communications Commission license needed for microwave transmission. However, costs were long an obstacle.

NASA went so far as to test lasers for transmitting signals between ground and space or between two spacecraft, but the results were discouraging. In December 1965, astronauts tried to send signals between the Gemini 6 and 7 spacecraft when they were simultaneously orbiting the Earth. They pointed a

[4]Jeff Hecht, *City of light: The story of fiber optics* (New York: Oxford University Press, 1999), p. 80.

[5]Jeff Hecht, *Beam: The race to make the laser* (New York: Oxford University Press, 2005).

hand-held transmitter, which contained four semiconductor diode lasers pulsed at 100 hertz to carry voice signals, between the two satellites. But the connection worked only briefly, probably because it was hard to aim the narrow beam at the other spacecraft. Later, NASA and the Air Force spent millions of dollars trying to develop high-speed laser links between satellites, but pointing and tracking proved insurmountable problems until recent years.[6]

With its primary interest in long-distance transmission, the telecommunications industry decided that the best approach was to develop an optical waveguide to carry laser signals. The logical approach seemed to be an optical version of the hollow metal waveguides similar to those used for microwave transmission—specifically the hollow circular guides that Bell Labs and others were developing to transmit frequencies around 60 gigahertz (GHz), called millimeter waves. Phone companies were running into the capacity limits of the chains of microwave towers that carried long-distance traffic at frequencies of a few gigahertz, so they were trying to move to higher frequencies. Millimeter waves were not transmitted well by the atmosphere, so phone companies planned to transmit them through buried waveguides. Bell was convinced that millimeter waveguides were the technology of tomorrow, but the parent AT&T was the country's monopoly carrier, so Bell had the luxury of planning for the day after tomorrow.

Metal pipes with reflective linings turned out to absorb too much light to transmit laser beams long distances. However, Bell Labs developed an ingenious scheme to repeatedly focus a laser beam through "gas lenses" formed periodically along the waveguide, so that the light would not touch the walls of the tube. It was a challenging and expensive system, but in theory it promised low loss, and Bell had plenty of time and research dollars.

1.2.1 Solid Optical Waveguides and Fiber Optics

Money was not as plentiful at Standard Telecommunications Laboratories (STL) in Harlow, England, although it was owned by the International Telephone and Telegraph conglomerate. STL was blessed with a visionary engineer heading its research programs—Alec Reeves—who in 1937 had invented pulse-code modulation, the basis of converting analog signals into digital form for transmission in modern networks. That invention had been so far ahead of its time that Reeves's patent had not earned a penny in royalties.

STL engineers experimented briefly with hollow optical waveguides, but the results were not encouraging, and so Reeves decided that STL should not pursue an expensive technology that was better suited to the wide open spaces of the United States than to smaller Britain. Instead, he turned his attention toward a

[6]Jeff Hecht, Reflections: Lasers as space-age technology, *Laser Focus World 30*, 8, pp. 45–47 (August 1994).

different type of microwave waveguide, flexible plastic rods known as dielectric waveguides. Their optical counterparts were fiber optics.

Fiber optics had originally been invented to transmit images from inaccessible places to the eye. The idea was to align many transparent fibers parallel to each other in a bundle, so that each one would essentially transmit one pixel of the image from one end to the other. The possibility of looking down the throat into the stomach intrigued physicians, and in 1930 Heinrich Lamm, a German medical student, assembled a short bundle and transmitted light through it. The image quality was not good because the fibers scratched each other and light leaked between them.

That problem was not solved until two decades later, when American optical physicist Brian O'Brien realized that he could trap light inside the fiber by covering it with a transparent cladding, making it a tiny optical waveguide. That invention opened the way to practical endoscopes for medical imaging, but nobody was thinking of communications because the most transparent glasses available had attenuation of one decibel per meter.

STL did not seek to duplicate those early optical fibers. Instead, Antoni Karbowiak set out to make an optical analog of microwave dielectric waveguides, which were solid plastic rods that guided microwaves along their exterior. Having worked on hollow millimeter waveguides, he sought to avoid one of their problems—propagation of the millimeter waves in multiple modes that could interfere with each other to generate noise. Karbowiak wanted an optical waveguide that would propagate light in only a single mode, but he found that would require a bare fiber only 0.1 to 0.2 micrometer (μm) in diameter, much too small for practical use. Then he left STL to accept a professorship in Australia.

Charles K. Kao, a young engineer born in Shanghai and trained in Britain, inherited the optical waveguide project. He had already been analyzing what would happen if the optical waveguide was clad with a layer of transparent material with lower refractive index. That cladding would confine the light within the fiber—the same conclusion O'Brien had reached a decade earlier. But Kao also found that if the difference between the refractive indexes of the core and the cladding was small, the core diameter could be increased to several micrometers and still transmit only a single mode. That larger core would collect light much more easily, and confine light much better, than a tiny bare fiber.[7]

Kao had essentially reinvented optical fibers, optimized for communications rather than for imaging. Bringing the guided light inside the fiber created a problem because the light had to go through the glass rather than air, which conventional wisdom held was inevitably more transparent. But Kao did not give up easily. Instead of asking how clear the best available glass was, he asked what

[7]Hecht, *City of Light,* Chapter 9.

was the fundamental lower limit on glass attenuation. Harold Rawson, a professor at the Sheffield Institute of Glass Technology in England, encouraged Kao with the information that impurities absorbed most of the light lost in standard glasses. If all the impurities could be removed, Rawson said, attenuation probably could be reduced below 20 dB/km, the target Kao had set to permit developing communication systems that carried telephone signals several kilometers between switching offices in adjacent communities.

With a younger colleague, George Hockham, Kao wrote a paper outlining their case for a single-mode fiber-optic communication system, which he presented at a January 27, 1966 meeting of the Institution of Electrical Engineers in London and later published in *Proceedings of the Institution of Electrical Engineers*.[8] They estimated that their system would have transmission capacity of a gigahertz, equivalent to nearly 200 analog video channels or 200,000 analog voice channels—more than was then available from coaxial cable or radio systems, and a huge increase over existing telephone trunk lines. The big problem was making a glass fiber as clear as they needed. Initial reactions were highly skeptical, and Bell Labs showed no interest. But Kao attracted the interest of two British government agencies—the defense ministry and the Post Office's telecommunications division.

Military contracts were a big part of STL's business, and the prospects for thin, flexible optical waveguides for use on the battlefield or inside military vehicles intrigued Don Williams of the Royal Signals Research and Development Establishment in Christchurch. Optical transmission promised a big advantage in the emerging world of electronic warfare. Electronic systems were vulnerable to jamming by enemy equipment and could be disabled by powerful bursts of electromagnetic energy from nuclear explosions. Optical transmission might present a way around those problems.

The Post Office Research Station, then at Dollis Hill in London, was hardly as stodgy as it sounds. It already was studying ideas for home phone customer access to remote computerized databases, a very early version of the Web. Critically, its research budget had just received a big boost. The Post Office also found Kao another important connection—the Corning Glass Works, a long-time leader in glass research.

The success of Kao's plan depended on removing impurities from glass, and that was a tough problem because ordinary glasses are made from inherently impure materials. However, Corning had earlier developed a technology for producing fused silica, which is essentially pure silicon dioxide. Corning physicist Robert Maurer saw two key drawbacks to using fused silica. Its extremely high melting point made fiber fabrication hard, and its refractive index was lower than

[8]K. C. Kao and G. A. Hockham, Dielectric-fiber surface waveguide for optical frequencies, *Proceedings IEE 113*, pp. 1151–1158 (July 1966).

other glasses, so something would have to be added to it to make the fiber core. But Maurer's gamble paid off. With Donald Keck, Peter Schultz, and Frank Zimar, Maurer managed to crack the 20 dB/km barrier in 1970.[9] He was surprised to find that no one else was even close.

The same year also saw another crucial development. Researchers at the Ioffe Physics Institute in Russia and Bell Labs in the United States demonstrated the first semiconductor diode lasers the could operate continuously at room temperature within weeks of each other. Their lasers lasted only minutes, but that marked tremendous progress on tiny lasers that were a perfect match for the tiny cores of optical fibers. Progress was also being made on LEDs, another potential light source.

1.2.2 Testing and Building Optical Systems

Engineers started testing systems long before they had low-attenuation fibers. In 1967, Richard Epworth, a young engineer just hired to work for Kao, used a laser to transmit black-and-white television signals through a 20-meter (m) bundle of high-loss fibers crossing a large voltage differential.[10] At about the same time, Northrop's Nortronics division demonstrated a battery-operated "fiber-optic data link" that transmitted 30-megahertz (MHz) signals from an LED through up to 7 m of bundled fibers to avoid electromagnetic interference and ground-loop problems.[11]

More demanding experiments soon followed. In late 1968, another young STL engineer, Martin Chown, demonstrated a 75-Mbit/s optical repeater using a diode laser sitting in a Dewar of liquid nitrogen.[12] By 1971 Chown and Murray Ramsay were able to transmit a strikingly clear color television signal through a small reel of fiber at the Centennial exhibition of the Institution of Electrical Engineers in London. It impressed Queen Elizabeth, and Lord Louis Mountbatten and Prince Philip stayed behind to ask Ramsay about the new system.[13] *Electronics*, then the field's leading trade magazine, highlighted fiber-optic progress in a feature.[14]

Bell Labs was slow to change course, thanks to its heavy investments in millimeter and hollow optical waveguides, as well as a reluctance to use outside ideas, called the "not invented here" syndrome. In mid-1970, a top engineering manager, Stew Miller, described a future in which fibers would be used for in-

[9]F. P. Kapron, D. B. Keck, and R. D. Maurer, Radiation losses in glass optical waveguides, *Applied Physics Letters 17*, pp. 423–425 (November 15, 1970).

[10]Hecht, *City of Light*, p. 123.

[11]Fiber optic data link assures interference-free signal transmission, *Laser Focus*, p. 18 (December 1967).

[12]Hecht, *City of Light*, p. 128.

[13]Ibid., p. 161.

[14]John N. Kessler, Fiber optics sharpens focus on laser communications, *Electronics*, pp. 46–52 (July 5, 1971).

teroffice trunks less than about 10 km, and confocal waveguides would span tens of kilometers without repeaters.[15] Corning's low-loss fiber changed those plans, but it took a couple of years before Bell quietly phased out the confocal waveguide program.

Meanwhile, the first primitive fiber-optic links started coming into use. The technology was neither cheap nor easy, the links were short, and the applications were in difficult environments where interference or voltage differentials made electronic transmission impossible. Mostly they transmitted data from measurement instruments.

1.2.3 Rapid Advances in Digital Communications

Ironically, the first applications of fiber optics in the computer industry were not in communications. Arrays of 12 optical fibers were used to illuminate the holes in punched cards during the 1960s. Computer uses at some major research universities and laboratories could access mainframes through remote terminals, but punched card input remained common into the early 1970s. ARPANET, the seed that would later become the Internet, had barely sprouted, linking only a handful of research sites.

Monopoly telephone carriers, led by the Bell System in the United States, defined the leading edge in telecommunications technology in the early 1970s. The public impression of industry innovation was dominated by Bell's Picturephone video-telephone system, which proved a dismal failure. But the crucial innovations reshaping the telephone system were deep inside the network. Starting in the 1960s, carriers had begun converting internal transmission from the traditional analog format to digital signals, using the pulse-code modulation system Reeves had invented.

The goal was to eventually convert all signals on the telephone network to digital form before multiplexing them for regional and long-distance transmission. Bell carefully planned the details, setting the standard for four levels of digital multiplexing. Copper wires could carry the two lowest speeds, the 1.5 Mbit/s T1 and the 6.3 Mbit/s T2 (originally developed to carry one Picturephone channel). The millimeter waveguide was expected to carry the highest speed, the 274-Mbit/s T4. Fiber appeared ideal to fit the middle 45 Mbit/s T3 level for trunk transmission between local telephone switches—just as Kao had proposed—filling an important gap.

However, Bell made a few changes to match its requirements. Worried about the problems of coupling light into a core only a few micrometers across, Bell shifted to multimode fibers with cores of 50 or 62.5 μm and a graded refractive

[15]Stewart E. Miller, Optical communications research progress, *Science 170*, pp. 685–695 (November 13, 1970).

index to increase bandwidth. That gave up the advantage of single-mode transmission, but Bell thought it would be good enough for 10- to 20-km links. For a laser source, Bell picked 850-nanometer (nm) gallium arsenide diode lasers, which were the most mature technology available.

All in all, it was an entirely reasonable design, which Bell put through exhaustive testing and field trials. The problem was that Bell management expected to phase the new fiber-optic equipment in over many years, as the telephone monopoly planned with the millimeter waveguide, which it had started developing in 1950.

Yet fiber technology did not stand still, making two key advances in short order. J. Jim Hsieh at Lincoln Labs developed a new family of semiconductor diode lasers based on InGaAsP, which emitted at wavelengths from 1.1 to 1.6 μm. And Masaharu Horiguchi at Nippon Telegraph and Telephone in Japan opened two new transmission windows in glass fibers, at 1.3 and 1.55 μm, with better transmission characteristics than at 850 nm.[16]

Lower attenuation at the longer wavelengths allowed transmission over longer distances. The new fibers also promised much higher bandwidth at 1.3 μm—but only in single-mode fibers. The new technology was a lifeline for the submarine cable group at Bell Labs, because their old coaxial cable technology could not keep up with satellite transmission. By 1980 they had begun developing the special-purpose technology needed for submarine fiber-optic cables, although the first transatlantic fiber cable was not laid until the end of 1988. But Bell management was not ready to give up on multimode fiber on land.

The critical push came from one of the upstart companies that had begun competing to carry long-distance traffic. MCI decided to upgrade its long-distance network by shifting to fiber optics and at the end of 1982 boldly bet on single-mode fibers transmitting 400 Mbit/s at 1.3 μm, because they could carry signals about 50 km between repeaters. Bell and other long-distance carriers followed, and soon single-mode fiber-optic networks spread across the country. Within a few years, data rates on the long-distance cables reached the gigabit range.

1.2.4 Fiber Optics for Data Communications

The fiber optics boom of the late 1970s and 1980s stimulated wide interest in the computer industry, which was turning to networking, minicomputers, and then microcomputers. Yet the ideas did not go far in the data communications world.

The fundamental issue was cost. Connecting a pair of computers with fiber required not just the fiber, but also a transmitter that converted electronic input to optical output, and a receiver that converted optical input to electronic output. It also needed expensive connectors that precisely aligned fibers to direct light

[16]Hecht, *City of Light*, Chapter 14.

between them. That cost much more than old-fashioned wires. The telecommunications industry could justify the expense because fibers could transmit signals at much higher speeds and over much longer distances than copper wires. Because data transmission did not need such high speeds or long distances, wires could almost always do the job.

There were a few exceptions. Military agencies developed special-purpose short fiber-optic links to meet requirements not encountered in the civilian world. As the Air Force began developing planes with airframes made of nonmetallic composite materials, engineers worried that on-board electronic systems would be vulnerable to electromagnetic interference, particularly from enemy electronic warfare equipment. That led to the installation of a 6-m fiber cable on the Marine Corps AV-8B Harrier jet to carry data from sensors to other equipment.

The Army developed a portable fiber-optic network to replace the 26-pair copper cables that had provided communication services in base camps. A promotional video compared a lightly built female soldier laying the fiber cable to two massive male soldiers hauling a heavy reel of the copper cable. The low attenuation of fiber cables made them ideal for linking field radar control centers to remote dishes; soldiers wanted the dish as far from the control center as possible in case an enemy missile homed in on the radar dish.

Fibers also found their way into some nonmilitary applications with difficult requirements. In 1988, in the first edition of my book *Understanding Fiber Optics*, I listed several reasons for using fiber-optic data links. The reasons included immunity to electromagnetic interference, better data security because fibers could not be tapped easily, the ability to make fiber cables nonconductive, and the elimination of spark hazards in environments such as oil refineries. An occasional factor was the small size of fiber-optic cables, which could allow easier installation. That meant that most fiber-optic data links were in a limited range of applications, generally in electromagnetically noisy environments, such as running alongside heavy power cables, or in secure environments where it was critical to prevent electromagnetic fields from leaking from a cable where they might be detected and used to decode the signal.

Another special application was in networks that required very high bandwidth for the time. The first standardized local area network (LAN) operating at 100 Mbit/s was the fiber distributed data interface (FDDI). Introduced in the mid-1980s, the original FDDI standard called for use of multimode graded-index fiber with either 62.5 or 85-μm cores and signal transmission using 1.3-μm light-emitting diodes (LEDs), which cost less than diode lasers and could transmit signals up to 2 km between nodes. Each node regenerated the output signal, and the entire network could contain up to 200 km of cable.[17] But at a time when

[17]Jeff Hecht, *Understanding fiber optics*, 1st ed. (Sams, Indianapolis, IN 1988).

1200-baud modems were standard for personal computers, few systems required 100 Mbit/s, and FDDI was not widely used. The companies trying to sell fiber-optic LANs could argue that installing fiber would provide room for future growth, but they did not succeed in selling many fiber-optic LANs. Nor did they expect the steady improvements in the bandwidth of copper cables, widely used in 100-Mbit/s Fast Ethernet, established as a standard in the mid-1990s.

1.2.5 The Internet and Fiber-Optic Booms

Fiber optics was in the right place at the right time to take advantage of the boom in competitive long-distance carriers in the 1980s, so fiber became the backbone of the new digital global telecommunication network. A new round of advances made it the right technology at the right time for the Internet boom of the 1990s.

The crucial advance was the invention of the erbium doped fiber amplifier, which grew from David Payne's research in specialty fibers at the University of Southampton in England. Doping the cores of optical fibers with rare-earth elements, Payne found that exciting the rare-earths with light from eternal lasers could make the rare-earth ions emit light. That soon led to making the light-emitting fibers work as lasers themselves. Payne's next step was to use the stimulated emission that oscillates in a laser to amplify an optical signal at the same wavelength in a fiber without a resonant cavity. He tested various rare-earth elements and concluded that the one best suited for optical amplification was erbium, which emits strongly across a range of wavelengths near 1.53 μm, close to the wavelength where standard optical fibers have their lowest loss. Early experiments in late 1987 recorded low noise and peak amplification of 26 decibels.[18]

Better yet, the experiments showed that erbium could amplify signals by at least 10 dB across a 25-nm range of wavelengths.[19] That broad range revived the idea of multiplexing signals at different wavelengths through the same fiber to multiply transmission capacity. Wavelength-division multiplexing was impractical in systems using electro-optical repeaters because the wavelengths had to be demultiplexed and put through separate repeaters. An optical amplifier with gain across a range of wavelengths could simultaneously amplify signals across the whole range.

Emmanuel Desurvire, then at Bell Labs, took a key step by showing that the optical amplifier could simultaneously amplify two 1-Gbit/s signals at separate

[18]R. J. Mears et al., High-gain rare-earth doped fiber amplifier at 1.54 μm, paper WI2 in *Technical Digest, Optical Fiber Communication Conference*, January 19–22, 1987, Reno, Nevada (Optical Society of America).

[19]R. J. Mears et al., Low-noise erbium-doped fiber amplifier operating at 1.54 μm, *Electronics Letters 23*, pp. 1026–1028 (September 10, 1987).

wavelengths without appreciable crosstalk.[20] Within months a race was on to see how many bits per second wavelength-division multiplexing could squeeze through a single fiber. In early 1990, a team from KDD in Japan sent 2.4-Gbit/s signals at four separate wavelengths through six erbium amplifiers and 459 km of fiber.[21] Others soon pushed to higher data rates by refining their technology, squeezing more optical channels closer together, modulating them at higher data rates, and adjusting fiber properties to increase transmission distances.

To use erbium optical amplifiers, developers had to shift system operation from the 1.3 μm of earlier systems to the 1.55-μm band where erbium amplified light. Fortuitously, attenuation of glass fibers is at its minimum in the erbium band, but chromatic dispersion of signals is much higher than at 1.3 μm. Therefore, fiber systems had to be redesigned to compensate for dispersion effects that otherwise would limit the maximum data rate. That took time, but laboratory data rates rose steadily, and in the mid-1990s wavelength-division multiplexed systems reached the market for long-distance transmission.

The timing could not have been better. Internet developers began routing Internet traffic through the global telecommunications network when total traffic was modest, and the Internet was little more than just another organization leasing lines to serve sites distributed around the United States. But Internet traffic began to take off with the dramatic expansion of the World Wide Web. The number of Web servers soared from 500 at the start of 1994 to 10,000 at the end of the year, with 10 million users.[22] For a brief interval in 1995 and 1996 Internet traffic doubled every three to four months as new users piled onto the Net.

Internet traffic never grew that fast again, despite myths that spread as the Internet boom evolved into a full-fledged bubble.[23] But it was clear to all that data traffic was growing much faster than the 10% a year growth of voice telephone traffic, and that handling that traffic would require expanding the capacity of the global telecommunications network. The timing was perfect for Ciena, Lucent, and Pirelli, which had introduced the first commercial wavelength-division multiplexing (WDM) systems in 1995 and 1996. WDM could deliver much more

---

[20]E. Desurvire, C. R Giles, and J. R. Simpson, Saturation-induced crosstalk in high-speed erbium-doped fiber amplifiers at $\lambda = 1.53$ μm, Paper TuG7 in *Technical Digest: Optical Fiber Communication Conference 1989*; Emmanuel Desurvire, C. Randy Giles, and Jay R. Simpson, Gain saturation effects in high-speed multichannel erbium-doped fiber amplifiers at $\lambda = 1.53$ μm, *Journal of Lightwave Technology 7*, pp. 2095–2104 (December 1989).

[21]H. Taga et al., 459 km, 2.4 Gbit/s 4 wavelength multiplexing optical fiber transmission experiment using 6 Er-doped fiber amplifiers, Postdeadline Paper 9, *Optical Fiber Communication Conference 1990*.

[22]http://public.web.cern.ch/Public/ACHIEVEMENTS/WEB/history.html as of August 31, 2007.

[23]K. G. Coffman and A. M. Odlyzko, Internet growth: Is there a "Moore's Law" for data traffic? in *Handbook of massive data sets*, J. Abello, P. M. Pardalos, and M. G. C. Resende, eds. (Kluwer, 2002), pp. 47–93, also available at http://www.dtc.umn.edu/~odlyzko/doc/networks.html

bandwidth per fiber, although it required installing different fiber than was used in most existing systems. Meanwhile, deregulation opened the telecommunications market up to new carriers, which raised money from investors eager to cash in on the sure thing of Internet growth, and expensive new fiber-optic systems were installed.

Equipment manufacturers poured money into research and development, and developers succeeded in squeezing more and more bits per second through fibers. In 1998 Bell Labs sent one hundred 10-Gbit/s channels through 400 km of fiber, a staggering 1 terabit per second,[24] and Lucent Technologies claimed it would have a commercial version transmitting at 40% of that rate available by the end of the year.[25]

By 2001, NEC Corporation and Alcatel had managed to push 10 terabits per second through single fibers, but by then the Internet bubble had burst. The telecommunications industry had run off the cliff, but the cartoon version of the law of gravity held the industry in midair briefly until it looked down and saw the ground was far below.[26] The bubble imploded with a visceral splat.

1.2.6 The Legacy of the Boom and Bust

The boom, the bubble, and the bust left the fiber optics industry with deep economic scars that are still healing. But the money pumped into the industry also left important technological legacies that are the foundation for modern fiber-optic data networks, from office LANs to the global telecommunications network.

The huge investment in fiber-optic cables and WDM systems left the global telecommunications network with much more transmission capacity than it needed, bringing down prices for both long-distance telephone calls and Internet traffic. Many installed fibers are still not carrying any traffic; few carry the maximum possible number of wavelength channels. In many areas there is plenty of room to expand transmission capacity without huge new cable installations, but traffic has filled cables in some areas.

Equipment manufacturers have developed cheaper versions of expensive bubble-era technologies. The dense WDM used to pack dozens of wavelength

---

[24]A. K. Srivastava et al., 1 Tbit/s transmission of 100 WDM 10 Gbit/s channels over 400 km of TrueWave fiber, Postdeadline paper PD 10, and S. Aisawa et al., Ultra-wide band, long distance WDM transmission demonstration: 1 Tbit/s (50×20 Bit/s0 600 km transmission using 1550 and 1580 nm wavelength bands, Postdeadline paper PD11, both at Optical Fiber Communication Conference, February 1998, San Jose (Optical Society of America, Washington, D.C.).

[25]Jeff Hecht, Planned super-Internet banks on wavelength-division multiplexing, *Laser Focus World 35* 5, pp. 103–105 (May 1998).

[26]Jeff Hecht, *City of light: The story of fiber optics, Revised and Expanded Edition* (Oxford University Press, 2004), Chapter 18.

channels on a single fiber required expensive laser transmitters and wavelength-separation optics. Increasing the separation from a fraction of a nanometer to 20 nm for coarse WDM has reduced costs so much that the technology can be used in high-speed local networks.

Optical component prices have come down dramatically as volumes have increased and manufacturing technology has improved. Costs are low enough that carriers around the world are installing fibers all the way to homes to provide premium broadband services, using single-mode fiber and coarse WDM.

Copper still dominates data transmission on the desktop and in homes. Copper is compatible with existing equipment and remains cheaper than fiber for transmitting high-speed signals over short distances or low-speed signals over longer distances. Verizon's fiber-to-the-home cables run parallel to its standard copper telephone wires in many communities. But fiber is playing an increasing role as data rates continue to soar. Gigabit Ethernet requires fiber for distances beyond 100 m, but special high-performance copper cable is needed to transmit 10-Gbit Ethernet more than 15 m.

This page intentionally left blank

2

Optical Fiber, Cables, and Connectors

Ulf L. Osterberg
Thayer School of Engineering, Dartmouth College. Hanover; New Hampshire 03755

2.1. LIGHT PROPAGATION

2.1.1. Rays and Electromagnetic Mode Theory

Light is most accurately described as a vectorial electromagnetic wave. Fortunately, this complex description of light is often not necessary for satisfactory treatment of many important engineering applications.

In the case of optical fibers used for tele- and data communication, it is sufficient to use a scalar wave approximation to describe light propagation in single-mode fibers and a ray approximation for light propagation in multimode fibers.

For the ray approximation to be valid, the diameter of the light beam has to be much larger than the wavelength. In the wave picture we will assume a harmonically time-varying wave propagating in the z direction with phase constant β. The electric field can be expressed as

$$E = E_0(x, y)\cos(\omega t - \beta z) \tag{2.1}$$

This is more conveniently expressed in the phasor formalism as

$$E = E_0(x, y)\, e^{j(\omega t - \beta z)} \tag{2.2}$$

where the real part of the right-hand side is assumed.

A wave's propagation in a medium is governed by the wave equation. For the particular wave in Eq. (2.2), the wave equation for the electric z component is

$$\nabla_t^2 E_z(x, y) + \beta_t^2 E_z(x, y) = 0 \tag{2.3}$$

where we have introduced

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

$$\nabla_t^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad \text{[TransverseLaplacian]}$$

$$\beta_t^2 = k^2 n^2 - \beta^2 \quad \text{[Transverse phase constant]}$$

$$k = \frac{2\pi}{\lambda} \quad \text{[Free space wave vector]}$$

$$n(x, y) \quad \text{[Refractive index]}$$

The variable kn corresponds to the phase constant for a plane wave propagating in a medium with refractive index n. There is an equivalent wave equation to Eq. (2.3) for the H_z component. We have to solve only the wave equation for the longitudinal components E_z and H_z. The reason for this is that E_x and E_y can both be calculated from E_z and H_z using Maxwell's equations.

2.1.2. Single-Mode Fiber

In an infinitely large isotropic and homogeneous medium, a light wave can propagate as a plane wave, and the phase constant for the plane wave can take on any value, limited only by the available frequencies of the light itself. When light is confined to a specific region in space, boundary conditions imposed on the light will restrict the phase constant β to a limited set of values. Each possible phase constant β represents a mode. In other words, when light is confined, it can propagate only in a limited number of ways.

For an engineer it is important to find out how many modes can propagate in the fiber, what their phase constants are, and their spatial transverse profile. To do this, we have to solve Eq. (2.3) for a typical fiber geometry (Fig. 2.1). Because of the inherent cylindrical geometry of an optical fiber, Eq. (2.3) is transformed into cylindrical coordinates and the modes of spatial dependence are described with the coordinates r, ϕ, and z. Because the solution is dependent on the specific refractive index profile, it has to be specified. In Fig. 2.2 the most common

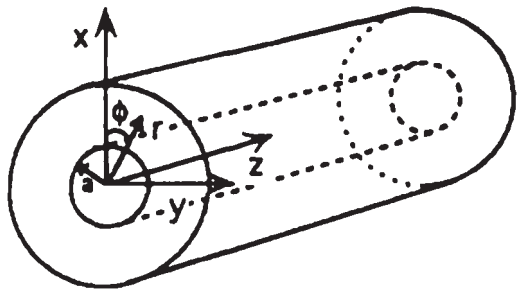


Figure 2.1 Typical fiber geometry. Reprinted from Ref. [1], p. 12, courtesy of Academic Press.

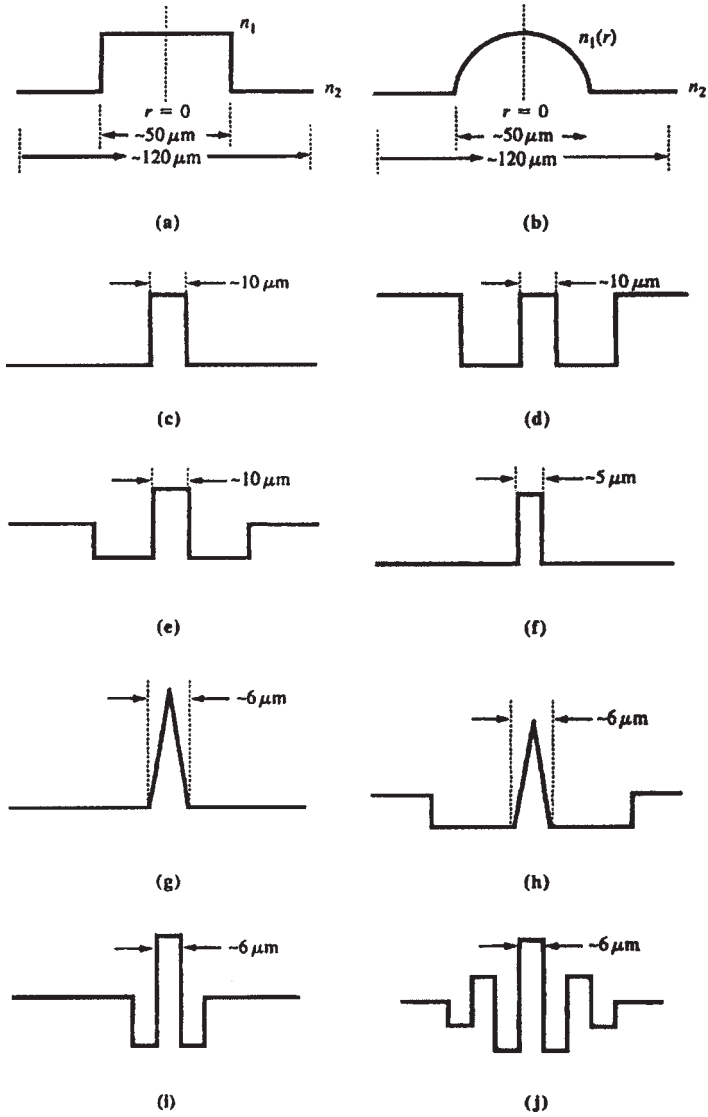


Figure 2.2 *Refractive* index profiles of (a) step-index multimode fibers, (b) graded-index multimode fibers, (c) match-cladding single-mode fibers, (d, e) depressed-cladding single-mode fibers, (f–h) dispersion-shifted fibers, and (i, j) dispersion-flattened fibers. Reprinted from Ref. [2], p. 125, courtesy of Irwin.

refractive index profiles are shown. For the step-index profile in Fig. 2.2c, a complete analytical set of solutions can be given [3]. These solutions can be grouped into three different types of modes: TE, TM, and hybrid modes, of which the hybrid modes are further separated into EH and HE modes. It turns out that for typical fibers used in tele- and data communication the refractive index difference between core and cladding, $n_1 - n_2$, is so small (~0.002–0.008) that most of the TE, TM, and hybrid modes are degenerate, and it is sufficient to use a single notation for all these modes—the LP notation. An LP mode is referred to as LP_{lm}, where the *l* and *m* subscripts are related to the number of radial and azimuthal zeros of a particular mode. The fundamental mode, and the only one propagating in a single-mode fiber, is the LP_{01} mode. This mode is shown in Fig. 2.3. To quickly figure out if a particular LP mode will propagate, it is useful to define two dimensionless parameters, V and b.

$$V = ka\sqrt{n_1^2 - n_2^2} \approx \frac{2\pi}{\lambda} a n_1 \sqrt{2\Delta} \tag{2.4}$$

where a is the core radius, λ is the wavelength of light, and $\Delta \cdot (n_1 - n_2)/n_1$. The V number is sometimes called the normalized frequency.

The normalized propagation constant b is defined as

$$b = \frac{\frac{\beta^2}{k^2} - n_2^2}{n_1^2 - n_2^2} \tag{2.5}$$

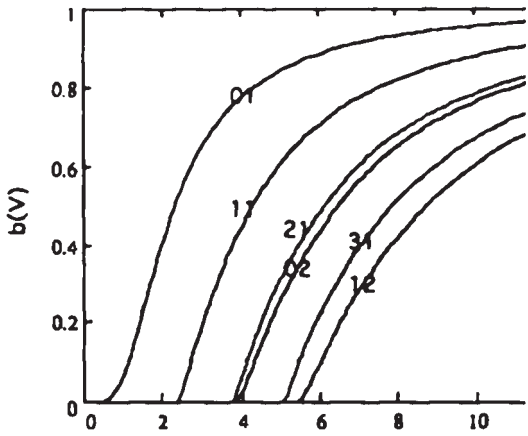


Figure 2.3 Cutoff frequencies for the lowest order LP modes. Reprinted from Ref. [1], p. 15, courtesy of Academic Press.

where b is the phase constant of the particular LP mode, k is the propagation constant in vacuum, and n_1 and n_2 are the core and cladding refractive indexes, respectively.

Equation (2.5) is very cumbersome to use because b has to be calculated from Eq. (2.3). For LP modes Marcuse et al. [4] have shown that to a very good accuracy the following formulas can be used to calculate b for different LP_{lm}, modes:

$$LP_{01}: b_{01} = 1 - \left[\frac{1+\sqrt{2}}{1+(1+V^4)^{\frac{1}{4}}}\right]^2 \tag{2.6}$$

$$LP_{lm}: b_{lm} = 1 - \frac{U_C^2}{V^2}\exp\left[\frac{2}{S}\left(\arcsin\left(\frac{S}{U_C}\right) - \arcsin\left(\frac{S}{V}\right)\right)\right]^2 \tag{2.7}$$

where

$$S = \sqrt{(U_C^2 - l^2 - 1)}$$

$$U_C = A - \frac{B-1}{8A} - \frac{4(B-1)(7B-31)}{3(8A)^3}$$

$$A = \pi\left[m + \frac{1}{2}(l-1) - \frac{1}{4}\right]; \quad B = 4(l-1)^2$$

The graphs in Fig. 2.4 were generated using Eqs. (2.6) and (2.7). The normalized propagation constant b can vary only between 0 and 1 for guided modes; this corresponds to

$$n_{2k} < \beta < n_1 k \tag{2.8}$$

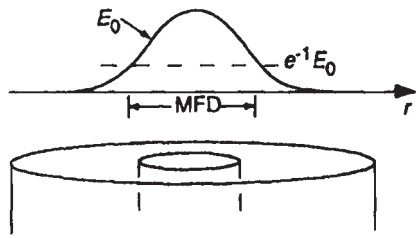


Figure 2.4 The electric field of the HE_{11} mode is transverse and approximately Gaussian. The mode field diameter is determined by the points where the power is down by e^{-2} or the amplitude is down by e^{-1}. The MFD is not necessarily the same dimension as the core. Reprinted from Ref. [6], p. 144, courtesy of Irwin.

Table 2.1
Cutoff Frequencies of Various $LP_{\ell m}$ Modes in a Step Index Fiber[a].

$\ell = 0$ modes	$J_1(V_c) = 0$	$\ell = 1$ modes	$J_0(V_c) = 0$
Mode	V_c	Mode	V_c
LP_{01}	0	LP_{11}	2.4048
LP_{02}	3.8317	LP_{12}	5.5201
LP_{03}	7.0156	LP_{13}	8.6537
LP_{04}	10.1735	LP_{14}	11.7915
$\ell = 2$ modes	$J_1(V_c) = 0$; $V_c \neq 0$	$\ell = 3$ modes	$J_2(V_c) = 0_i$; $V_c \neq 0$
LP_{21}	3.8317	LP_{31}	5.1356
LP_{22}	7.0156	LP_{32}	8.4172
LP_{23}	10.1735	LP_{33}	11.6198
LP_{24}	13.3237	LP_{34}	14.7960

[a]Reprinted from Ref. [5], p. 380, courtesy of Cambridge University Press.

The wavelength for which b is zero is called the cutoff wavelength; that is,

$$b_{lm}(V_{CO}) = 0 \Rightarrow \lambda_{CO} = \frac{2\pi}{V_{CO}} a n_1 \sqrt{2\Delta} \tag{2.9}$$

Therefore, for wavelengths longer than the cutoff wavelength, the mode cannot propagate in the optical fiber.

Cutoff values for the V number for a few LP modes are given in Table 2.1. The fundamental mode can, to better than 96% accuracy, be described using a Gaussian function

$$E(r) = E_0 \exp\left[-\left(\frac{r}{w_g}\right)^2\right] \tag{2.10}$$

where E_0 is the amplitude and $2w_g$ is the mode field diameter (MFD) (Fig. 2.4). The meaning of the MFD is shown in Fig. 2.5. The MFD for the fundamental mode is larger than the geometrical diameter in a single-mode (SM) fiber and much smaller than the geometrical diameter in a multimode (MM) fiber. The optimum MFD is given by the following formula [7]:

$$\frac{w_g}{a} = 0.65 + 1.619V^{\frac{-S}{2}} + 2.87V^{-6} \tag{2.11}$$

where a is the core radius. Equation (2.11) is valid for wavelengths between $0.8\,\lambda_{CO}$ and $2\,\lambda_{CO}$.

If the radial distribution for higher order modes is needed, it is necessary to use the Bessel functions [3]. In Fig. 2.6 the radial intensity distribution is shown

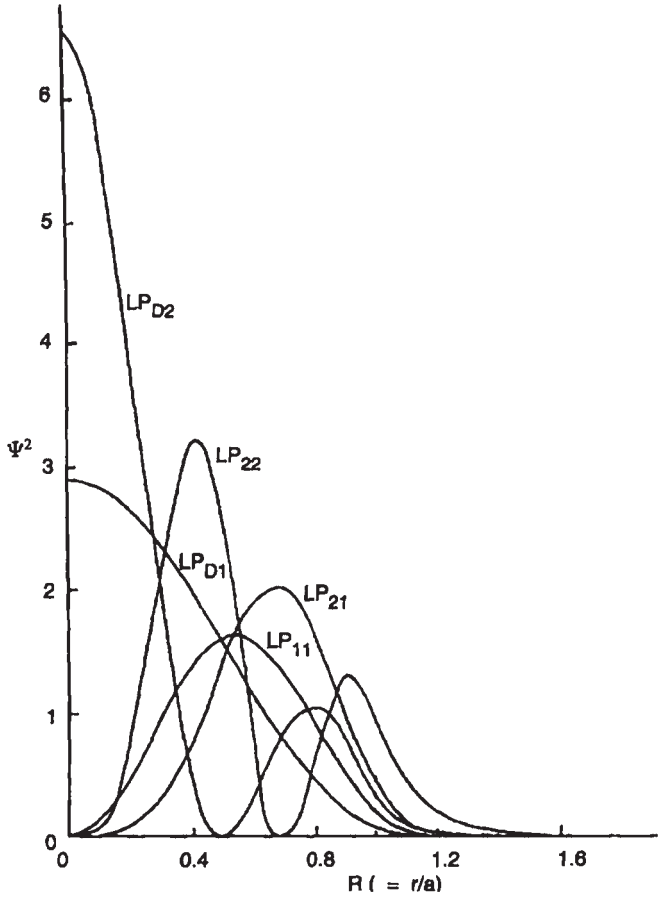


Figure 2.5 Radial intensity distributions (normalized to the same power) of some low-order modes in a step-index fiber for V = 8. Notice that the higher order modes have a greater fraction of power in the cladding. Reprinted from Ref. [5], p. 382, courtesy of Cambridge University Press.

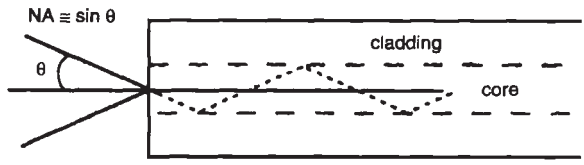


Figure 2.6 Acceptance angle for an optical fiber. Reprinted from Ref. [1], p. 10, courtesy of Academic Press.

Table 2.2
CCITT Recommendation G.652[a].

Parameters	Specifications
Cladding diameter	125 μm
Mode field diameter	9–10 μm
Cutoff wavelength λ_{co}	1100–1280 nm
1550-nm bend loss	≤1 dB for 100 turns of 7.5 cm diameter
Dispersion	≤3.5 ps/nm · km between 1285 and 1330 nm ≤6 ps/nm · km between 1270 and 1340 nm ≤20 ps/nm · km at 1550 nm
Dispersion slope	≤0.095 ps/nm^2 · km

[a]Reprinted from Ref. [2]. p. 126, courtesy of Irwin.

for five LP modes in a fiber with V = 8. Recommended specifications for a single-mode fiber are summarized in Table 2.2.

2.1.3. Multimode Fiber

The previous discussion has in principle been for a step-index MM fiber. Because of the severe differences in propagation time between different modes in a step-index fiber, these are not commonly used in practice. Instead, a graded refractive index core is used for a MM fiber.

The various graded-index profiles are generated by

$$n^2(r) = \begin{cases} n_1^2\left(1 - 2\Delta\left(\dfrac{r}{a}\right)^q\right); 0 \le r \le a \\ n_1^2(1 - 2\Delta); r \ge a \end{cases} \tag{2.12}$$

For different q's, q is called the profile exponent. The optimum profile is the one that gives the minimum dispersion; this occurs for q at slightly less than 2.

The total number of modes that can propagate in a MM fiber is given by

$$N = \frac{1}{2}\frac{q}{q+2}V^2 \tag{2.13}$$

For a step-index fiber q = ∞ and N = $V^2/2$. Equation (2.13) is only valid for large V numbers.

One can calculate the different ray paths that are possible in a MM fiber using a geometrical optics approach [8]. A more accurate view of the different modes can be obtained by solving Eq. (2.3) using the so-called WKB (Wigner-Kramer-

Brillouin) approximation [3]. Using this analysis, we can show that the phase constants for the different modes obey the following relationship:

$$\beta_m = nk\sqrt{1-2\Delta\left(\frac{m}{n}\right)^{\frac{q}{q+2}}}, \quad m = 1, 2, \ldots, N \tag{2.14}$$

2.1.4. Optical Coupling

A first approach to estimate how much light can be coupled into an optical fiber is to use the ray picture. In this picture, the light is confined within the core if it undergoes total internal reflection at the core-cladding boundary. This will occur only for light entering the fiber within an acceptance cone defined by the angle θ (Fig. 2.6). Rather than stating the angle θ for an optical fiber, it is the convention to give sin θ, which is called the numerical aperture (NA). The NA is defined as

$$NA = n\sin\theta \tag{2.15}$$

where n is the refractive index of the medium from where the light is coming. In the case of coupling into an optical fiber, light is usually coming from air and subsequently $n < 1$. A more useful formula for the NA can be obtained if we use the dimensionless parameter △,

$$NA \approx n_1\sqrt{2\Delta} \tag{2.16}$$

For an incoherent light source such as a light-emitting diode (LED), one can show that the total power accepted by the fiber is given by [9]

$$P \approx \pi B A_{fiber} (NA)^2 \tag{2.17}$$

where B is the LED's radiance (units for radiance is watts per area and steradian).

It is more common to give a coupling efficiency; thus, giving the total power accepted by the fiber, the efficiency is defined as [10]

$$\eta = \frac{P_{lm}}{P_{in}} \tag{2.18}$$

where P_{in} is the power launched into the fiber and P_{lm} is the power accepted by the LP_{lm} mode. For link budget analyses it is more convenient to deal with coupling losses in units of decibels α:

$$\alpha = 10\log\eta \tag{2.19}$$

A more general definition of the coupling efficiency η for an optical fiber that obeys the weekly guiding approximation can be given as [10]

$$\eta = \left| \frac{n_2}{z_0} \int_A \int_\infty E_i E_{lm}^* dA \right|^2 \tag{2.20}$$

where n_2 is the refractive index of the cladding, z_o is the free space wave impedance, and E_i and E_{lm} are the electric field amplitudes for the incoming light and for the light propagating in mode LP_{lm} in the fiber, respectively.

Coherent light from a laser can often be approximated with a Gaussian beam. Furthermore, if we restrict ourselves to a SM fiber, so that the LP_{01} mode can also be approximated as a Gaussian field, it is possible to calculate η analytically [7, 11]

$$\eta = \left(\frac{4D}{B} \right) e^{\frac{-AC}{B}} \tag{2.21}$$

where

$$A = \frac{(k/s_1)^2}{2}$$

$$B = G^2 + (D + 1)^2$$

$$C = (D + 1)F^2 + 2DFG\sin\theta + D(G^2 + D + 1)\sin^2\theta$$

$$D = \left(\frac{s_2}{s_1} \right)^2$$

$$F = \frac{2\Delta}{k/s_1^2}$$

$$G = \frac{2\Delta Z}{k/s_1^2}$$

$$k' = \frac{2\pi n_0}{\lambda}$$

where $\triangle Z$ is the separation distance between source and fiber, $\triangle$ is the lateral displacement, θ is the angular displacement, and s_1 and s_2 are the mode field radii or spot sizes of the source field and fiber field, respectively. The refractive index of the medium between source and fiber is denoted n_0.

Equation (2.21) takes into account four different coupling cases at once (Fig. 2.7). If only one of these different coupling cases is present at a time, Eq. (2.21) can be simplified to:

Case 2.1. Spot-size mismatch s_1 not equal to s_2:

$$\eta = \left(\frac{2s_2 s_2}{s_1^2 + s_2^2} \right)^2 \tag{2.22}$$

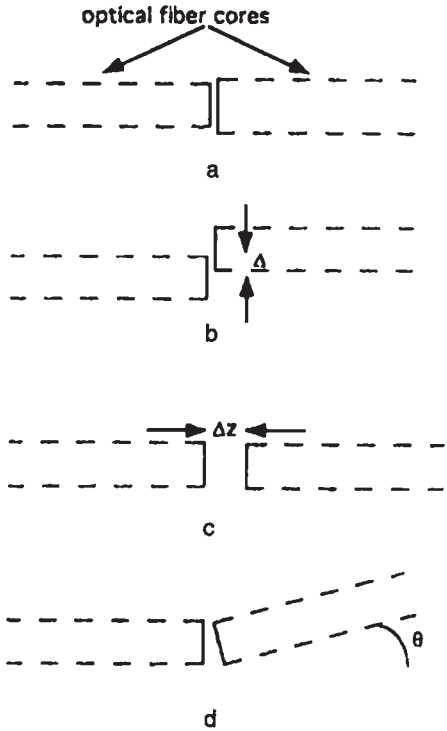


Figure 2.7 Different coupling cases. Reprinted from Ref. [1], p. 18, courtesy of Academic Press.

Case 2.2. Transverse offset △:

$$\eta = \exp\left(-\left(\frac{\Delta}{s_2}\right)^2\right) \tag{2.23}$$

Case 2.3. Longitudinal offset △Z:

$$\eta = \frac{1}{1+\left(\frac{\Delta Z}{zZ_0}\right)^2} \tag{2.24}$$

where Z_0 is the Rayleigh range.

Case 2.4. Angular misalignment θ:

$$\eta = \exp\left(-\left(\frac{\theta}{\theta_0}\right)^2\right) \tag{2.25}$$

If a lens is used in between the emitter and fiber, some modifications to the previous formulas have to be done. What the lens can do for us is to match the output angle of the emitter to the acceptance angle of the receiving fiber. If properly done, the power coupled into the fiber is multiplied with the lens magnification factor M: $M = d_{rec}/d_{em}$. All the preceding formulas need to be corrected for reflection losses. If the refractive index of the medium between the source and the fiber is denoted n_0, the coupled power into the fiber is reduced with a factor R,

$$R = \left(\frac{n_1 - n_0}{n_1 - n_0} \right)^2 \tag{2.26}$$

2.2. OPTICAL FIBER CHARACTERIZATION

2.2.1. Optical Fiber Materials

The material is primarily chosen to provide minimum attenuation. Table 2.3 shows order-of-magnitude attenuation at three different wavelengths for four common glass types. An introduction to optical fiber fabrication techniques and optimization of different fiber types for bandwidth or other properties is given in Chapter 4.

For tele- and data communication fibers, fused silica glass is the preferred material. To provide guiding of the light, the core of the fiber is doped with a small molar percentage of a substance that increases the refractive index. It is also possible to dope the cladding so that its refractive index becomes lower than the pure silica glass index in the core.

2.2.2. Attenuation

Attenuation is a very important factor in designing effective long-distance fiber-optic networks. Consequently, the fabrication methods have improved

Table 2.3

Scattering Loss for Several Representative Glass Materials[a].

	dB/km		
Material	633 nm	800 nm	1060 nm
Fused silica	4.8	1.9	0.6
Soda lime	8.5	3.3	1.1
Borosilicate crown	7.7	3.0	1.0
Lead silicate	47.5	18.6	6.0

[a]Reprinted from Ref. [1], p. 21, courtesy of Academic Press.

Table 2.4
Factors Affecting Attenuation[a].

Intrinsic loss mechanisms
Tail of infrared absorption by Si–O coupling
Tail of ultraviolet absorption due to electron transitions in defects
Rayleigh scattering due to spatial fluctuations of the refractive index
Absorption by impurities
Absorption by molecular vibration of OH
Absorption by transition metals
Structural imperfections
Geometrical nonuniformity at core–cladding boundary
Imperfection at connection or splicing between fibers

[a]From Ref. [3].

dramatically during the past 30 years so that attenuation is measured in a few tenths of a dB/km. The dB is defined in Eq. (2.19). The various factors affecting the attenuation, in the 0.8- to 1.6-μm wavelength region, are listed in Table 2.4. Figure 2.8 shows schematically how some of the factors contributing to the overall light attenuation vary with the wavelength in the near-infrared wavelength region. Typical total losses for an optical fiber, at the three different transmission windows at 800, 1300, and 1550 nm, are <2 or 3, <0.5, and <0.2 dB/km, respectively. These numbers are for SM fibers; from Fig. 2.9 it can be seen that MM fibers have slightly higher losses. Additional information on spectral attenuation of optical fibers is provided in Chapter 11 and Chapter 15. The losses dealt with to date have been due to either intrinsic properties of the glass or extrinsic properties (such as OH and transition metal contents) that come from the particular fabrication method used. In addition, there are bending losses. If the fiber has been improperly cabled or installed, these bending losses can be substantial. Bending losses are divided into micro- and macrobending losses. Micro-bending losses are due to nanometer-size deviations in the fiber, whereas macro-bending losses are due to visible bends in the fiber. Figure 2.10 shows qualitatively how micro- and macrobends contribute to the overall loss in a SM and MM fiber.

2.2.3. Dispersion

Dispersion occurs because different wavelengths experience different propagation constants, $\beta(\lambda)$, and therefore travel with different velocities causing a longer temporal pulse at the end of the fiber. Dispersion does not alter the wavelength (frequency) content of the light pulse. From a communications point of view, dispersion is a very important factor because it directly affects the bit rate. The following are three major components contributing to the dispersion:

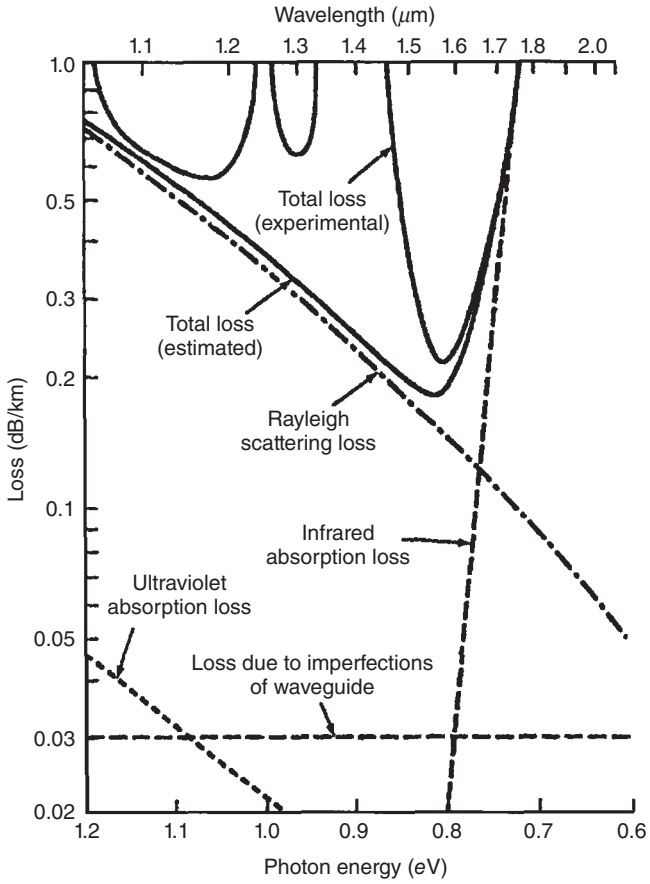


Figure 2.8 Transmission loss in silica-based fibers. Reprinted from Ref. [12], p. 474, courtesy of Irwin.

- Material
- Waveguide
- Intermodal

The material and waveguide dispersion are often referred to as chromatic dispersion, measured in units ps/nm-km, which is defined as

$$D_{chr} = -\frac{1}{L}\frac{dt_g}{d\lambda} \tag{2.27}$$

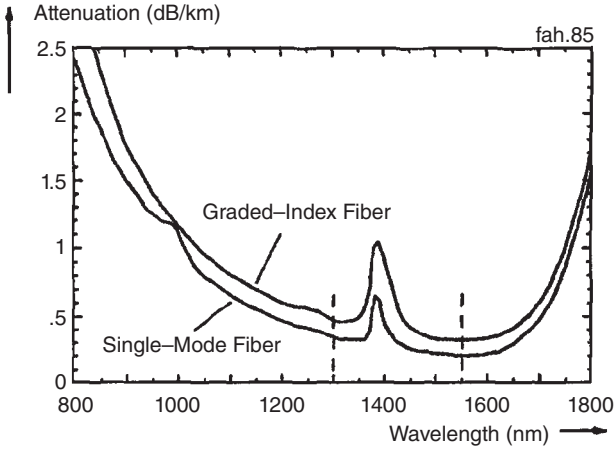


Figure 2.9 Wavelength dependence of fiber attenuation. Reprinted from Ref. [14], p. 8. Copyright 1989 by Hewlett-Packard Company. Reproduced with permission.

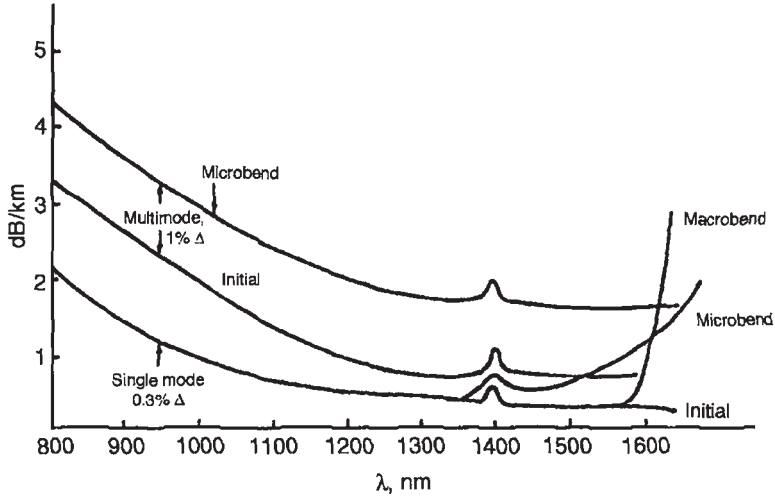


Figure 2.10 Bend-induced losses of optical fibers. Reprinted from Ref. [15], p. 1.33, courtesy of McGraw-Hill.

where L is the fiber length and t_g is the time required to propagate a distance L. The subscript g refers to group velocity. When a pulse propagates in a dispersive medium, it propagates with the group velocity,

$$V_g = \frac{d\omega}{d\beta} \tag{2.28}$$

The phase travels with the phase velocity given by

$$V_p = \frac{\omega}{\beta} = \frac{c}{n} \tag{2.29}$$

The dispersion properties are completely determined by the group velocity.

Material dispersion originates from the fact that the refractive index is a function of wavelength, $\eta(\lambda)$:

$$D_{mat} = -\frac{\lambda}{c}\frac{d^2 n_1}{d\lambda^2} \tag{2.30}$$

The refractive index can be calculated from the Sellmeier equation:

$$n_1^2(\lambda) = 1 + \sum_{i=1}^{3} \frac{a_i \lambda^2}{\lambda^2 - b_i} \tag{2.31}$$

where the constants a_i and $_i$ may vary with the particular fiber composition. These constants have been accurately measured and tabulated for different Ge, B, and P concentrations [16].

Waveguide dispersion is due to the different propagation characteristics of the light in the core and cladding. Keep in mind that a large portion of the light (30–40%) travels in the cladding for the LP_{01} mode around the cutoff frequency for the next higher order mode. The waveguide dispersion can be approximated, using the normalized propagation constant b, as

$$D_{wg} \approx -\frac{\lambda}{c} n_1 \Delta \frac{d^2 b}{d\lambda^2} \tag{2.32}$$

The total chromatic dispersion as well as its constituents are plotted in Fig. 2.11. As shown in Fig. 2.2, an optical fiber can have many different refractive index profiles. How the refractive index profile can affect the dispersion properties of the fiber is shown in Fig. 2.12.

Intermodal dispersion arises from the different travel times for the different modes in the fiber. For modal dispersion the units are nsec/km, and D_{mod} is defined as follows [2]:

$$D_{mod} = t_g^{max} - t_g^{min} \tag{2.33}$$

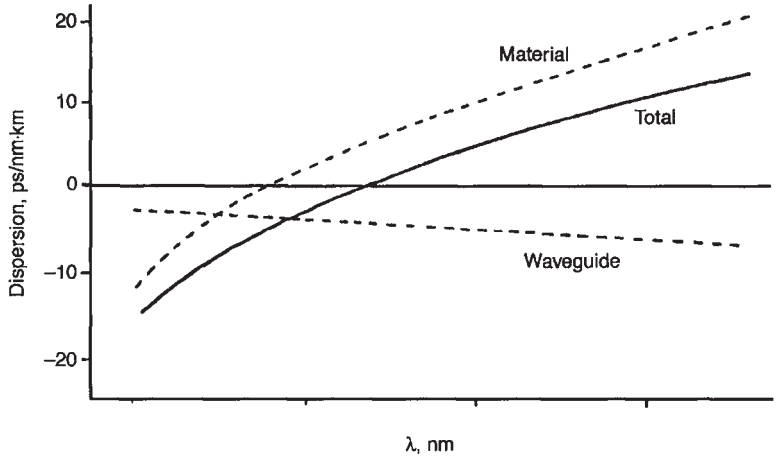


Figure 2.11 Dispersion vs. wavelength for a single-mode fiber. Reprinted from Ref. [15], p. 1.38, courtesy of McGraw-Hill.

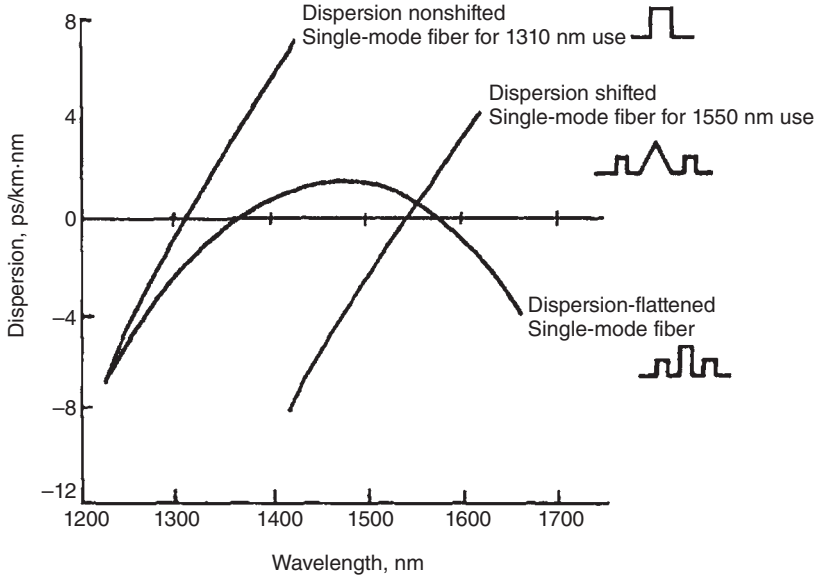


Figure 2.12 Spectral disposition for various single-mode fiber profiles. Reprinted from Ref. [15], p. 138, courtesy of McGraw-Hill.

where t_g^{max} and t_g^{min} are the propagation times for the modes traveling the longest and shortest distances, respectively. Using the group index N_1 [3],

$$N_1 = n_1 + k\frac{dn_1}{dk} \tag{2.34}$$

where the term $k\, dn_1/dk$ is the material dispersion for a plane wave traveling in the core region; we can approximately write the modal dispersion in a step-index fiber as follows [2]:

$$D_{mod} \cong -\frac{N_1}{c}\Delta \tag{2.35}$$

and for a graded-index fiber, with $q = 2(1 - \triangle)$, as

$$D_{mod} \cong -\frac{N_1}{c}\frac{\Delta^2}{8} \tag{2.36}$$

Note that for most fibers $N_1 < n_1$ is an excellent approximation [3].

The total fiber disposition in a MM fiber, units nsec/km, can be shown to be the following [2]:

$$D_{tot}^2 = D_{chr}^2 \Delta\lambda^2 + D_{mod}^2 \tag{2.37}$$

where $\triangle\lambda$ is the spectral bandwidth of the light source.

Knowledge of the total dispersion, D_{tot}, can be transmitted into bandwidth (BW), units MHz km, as follows [17]:

$$BW \cong \frac{350}{D_{tot}} \tag{2.38}$$

2.2.4. Mechanical Properties

Optical fibers have outstanding signal characteristics such as low attenuation and small dispersion. For optical fibers to be practically useful they must also have good mechanical properties to withstand environmental strains.

Silica glass has a very high threshold for breaking when there are no flaws, either on the surface or inside the glass. The tensile strength has been estimated to be 20 GPa [19]. Because optical glass fibers have a large surface-to-volume ratio, they are very prone to failure from surface flaws that propagate into the glass. When fibers are tested for their mechanical reliability, it is common to use short fibers for destructively testing the fibers' breaking threshold for static and dynamic loads. In addition, a screen test is performed on long samples of the fibers to find the effective strength [19].

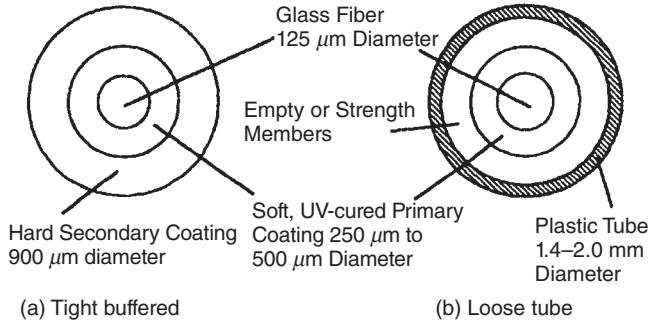


Figure 2.13 Common single-fiber cable types. Reprinted from Ref. [1], p. 36, courtesy of Academic Press.

2.2.5. Cable Designs

Because an optical fiber is very brittle and susceptible to chemical degradation, it is important to protect the optical performance. Two common types of fiber-optic cable construction are shown in Fig. 2.13. Common for all cable types is that the individual fibers are coated with some kind of organic material, examples of which include the following [16]:

- Polydimethyl siloxane
- Silicone oils
- Extrudates
- Acrylates

For ease of manufacturing and to prevent the glass surface from being scratched, the coating is applied during the drawing process. Often, the first coating is not sufficient to protect the fiber, and so additional coatings are used, which are commonly referred to as buffering. The buffered fiber is then embedded in various strength members and additional protective jackets. The primary goals of these additional layors is to protect the fibers from water, rodents, and temperature fluctuations and to reduce stress in the fiber (Fig. 2.14).

2.2.6. Connectors

It is important to be able to easily reconfigure communication links for cost and performance optimization, to replace link subunits, and to relocate equipment [1]. To do this, connectors are a key component in the vast majority of optical links (a recent exception is the introduction of active optical cables that permanently connectorize the transceiver assembly, as described in Chapter 4). The two types of fiber-optic connections are:

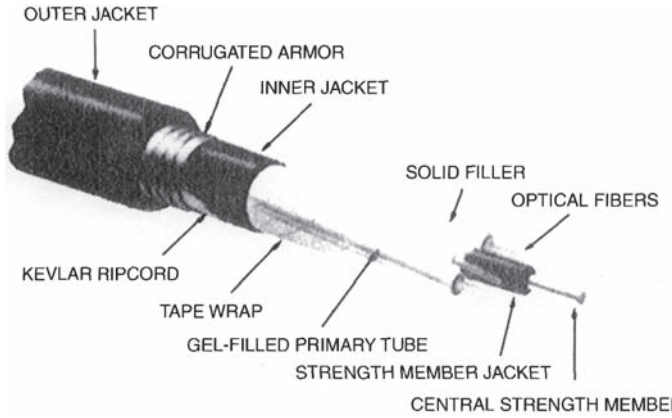


Figure 2.14 Fiber-optic cable-corrugated armor. Reprinted from Ref. [22], p. 247, courtesy of Academic Press.

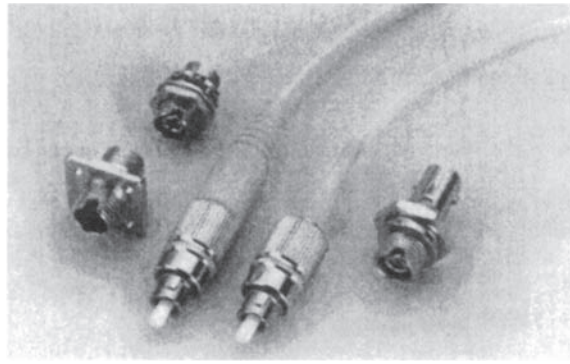


Figure 2.15 FC connectors and adaptors. Reprinted from Ref. [24], p. 16, courtesy of Molex Fiber Optics, Inc.

- Fiber to transceiver
- Fiber to fiber

The fiber-to-fiber connectors are divided into simplex (one fiber to one fiber) and duplex (two fibers to two fibers) connectors. The two basic connector designs are butt joints and expanded beam joints.

Connectors are further differentiated depending on the particular mechanical method used to join the connectors.

FC connectors use threaded fasteners (Fig. 2.15), and SC connectors use a spring latch (Fig. 2.16). These connectors can come as both simplex and duplex.

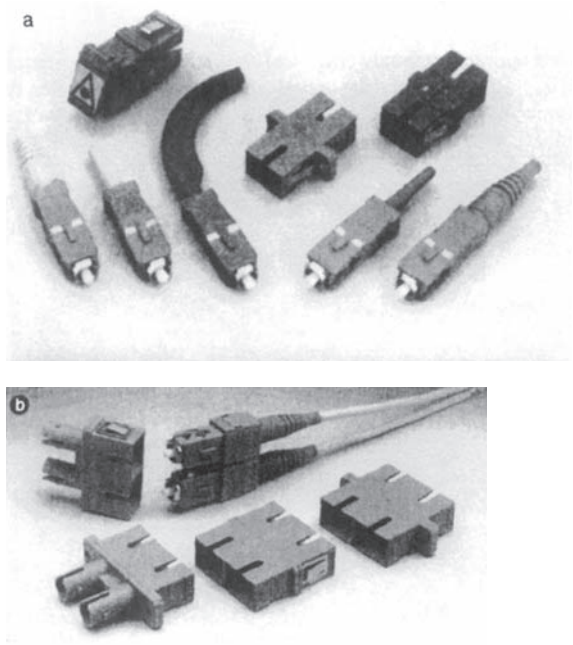


Figure 2.16 (a, b) Spring Latches. Reprinted from Ref. [24], p. 24 (a), and p. 34 (b), courtesy of Molex Fiber Optics, Inc.

Variations on the SC connector, developed for specific standardized local area networks, are FDDI, ESCON, and FCS compliant or SC Duplex connectors (Figs. 2.17a–c). Some early versions of the FCS-compliant connector featured keying mechanisms located in different orientations than those used today, or with different width keys for single-mode and multimode applications; these variants are no longer in common use. Many types of older connectors can still be found in the field, such as the multimode SMA connector (Fig. 2.18). More recent installations will make use of small form factor connectors; these are discussed in detail in Chapter 3.

Obviously, some additional loss is introduced into a communications link with the incorporation of a connector. Both a local and a distributed loss result from the connector joint. The local loss factors can be calculated using the formulas in Section 2.1.4. The distributed losses are due to changes in the model launching conditions [15].

Connector losses can be as high as 0.5 dB (this corresponds to an additional km of fiber at 1.3 pm). With improved manufacturing technology, losses between 0.1 and 0.2 dB are obtained for new connectors. Some typical connection loss values useful in link budget planning are given in Chapter 12.

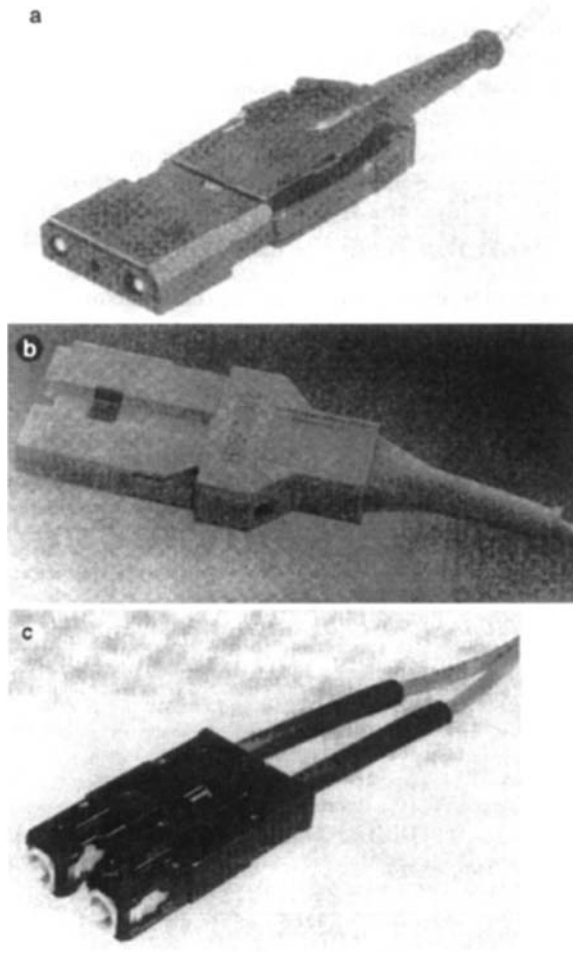


Figure 2.17 Duplex connectors: (a) ESCON (R), (b) FDDI, and (c) FCS duplex (fiber channel standard compliant). Reprinted from Ref. [1], p. 50, courtesy of Academic Press.

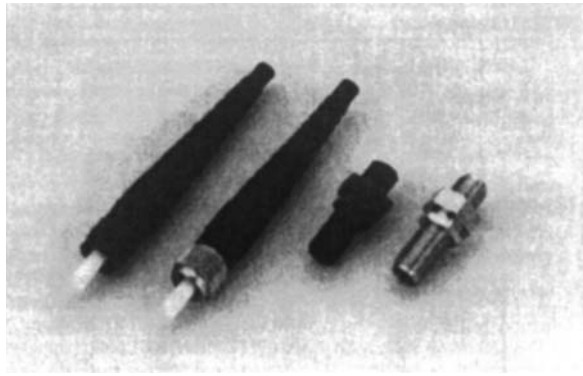


Figure 2.18 SMA connectors and adapters. Reprinted from Ref. [24], p. 42, courtesy of Molex Fiber Optics, Inc.

REFERENCES

1. Webb, J. R., and U. L. Osterberg. 1995. In *Optoelectronics for Datacommunication.* San Diego: Academic Press.
2. Liu, M. M. K. 1996. Fiber, cable and coupling. In *Principles and Applications of Optical Communications*, eds. R. C. Lasky, U. L. Osterberg, and D. P. Stigliani. Chicago: Irwin.
3. Okoshi, T. 1982. *Optical Fibers*. New York: Academic Press.
4. Marcuse, D., D. Gloge, and E. A. J. Marcatiti. 1979. *Guiding Properties of Fibers: Optical Fiber Telecommunications*. Orlando, Fl.: Academic Press.
5. Ghatak, A., and K. Thyagarajan. 1989. *Optical Electronics*. Cambridge: Cambridge University Press.
6. Pollock, C. R. 1995. *Fundamentals of Optoelectronics*. Chicago: Irwin.
7. Marcuse, D. 1977. Loss analysis of single-mode finger splices. *Bell System Tech. J.* 56:703–718.
8. Cheo, P. K. 1990. *Fiber Optics and Optoelectronics.* 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
9. Powers, J. P. 1993. *An Introduction to Fiber Optic Systems.* Homewood, Ill.: Aksen.
10. Neumann, E. G. 1988. *Single-Mode Fibers: Fundamentals*. Berlin: Springer-Verlag.
11. Nemoto, S., and T. Makimoto. 1979. Analysis of splice loss in single-mode fibers using a Gaussian field approximation. *Opt. Quantum Electron.* 11:447–457.
12. Maurer, R. D. 1973. Glass fibers for optical communication. *Proc. IEEE* 61:452–462.
13. Chen, C. L. 1996. *Principles and Applications of Optical Communications.* Chicago: Irwin.
14. Hentschel, C. 1989. *Fiber Optics Handbook.* Boeblingen, Germany: Hewlett-Packard.
15. Allard, F. C. 1990. *Fiber Optics Handbook for Engineers and Scientists.* New York: McGraw-Hill.
16. Shibata, N., and T. Edahiro. 1982. *Trans. Inst. Electron. Commun. Eng. Jpn.* 65:166–172.
17. Hecht, J. 1987. *Understanding Fiber Optics.* Indianapolis, Ind.: Sams.
18. Murata, H. 1996. *Handbook of Optical Fibers and Cables*. New York: Dekker.
19. Kalish, D., P. L. Key, C. R. Kurkjian, B. K. Tanyal, and T. T. Wang. 1979. Fiber characterization-mechanical. In *Optical Fiber Telecommunications*, eds. S. E. Miller and C. G. Chynoweth. Boston: Academic Press.

20. Philen, D. L., and W. T. Anderson. 1988. Optical fiber transmission evaluation. In *Optical Fiber Telecommunications II*, eds. S. E. Miller and I. P. Kaminow. Boston: Academic Press.
21. Gartside, C. H., P. D. Patel, and M. R. Santana. 1988. Optical fiber cables. In *Optical Fiber Telecommunications II*, eds. S. E. Miller and I. P. Kaminow. Boston: Academic Press.
22. Radcliffe, J., C. Paddock, and R. Lasky. 1995. Integrated circuits, transceiver modules and packaging. In *Optoelectronics for Datacommunication*, eds. R. C. Lasky, U. L. Osterberg, and D. P. Stigliani. San Diego: Academic Press.
23. Young, W. C., and D. R. Frey. 1988. Fiber connectors. In *Optical Fiber Telecommunications II*, eds. S. E. Miller and I. P. Kaminow. Boston: Academic Press.
24. Molex Fiber Optics, Inc. 1996. *Fiber Optic Product Catalog*, No. 1096.
25. Hill, K. O., Y. Fujii, and D. C. Johnson, et al. 1978. Photosensitivity in optical fiber Waveguides: Application to reflection filter fabrication, *Applied Physics Letters* 32(10):647–649.

3

Small Form Factor Fiber-Optic Interfaces

John Fox
ComputerCrafts Inc., Hawthorne, N.J. 07507

Casimer DeCusatis
IBM Corporation, Poughkeepsie, N.Y. 12601

3.1. INTRODUCTION

Conventional duplex fiber-optic connectors, such as the SC Duplex defined by the ANSI Fibre Channel Standard [1], achieve the required alignment tolerances by threading each optical fiber through a precision ceramic ferrule. The ferrules have an outer diameter of 2.5 mm, and the resulting fiber-to-fiber spacing (or pitch) of a duplex connector is approximately 12.5 mm. Since the outer diameter of an optical fiber is only 125 μm, it should be possible to design a significantly smaller optical connector. Smaller connectors with fewer precision parts could dramatically reduce manufacturing costs and have the potential to open up new applications such as fiber to the desktop. Smaller connectors and transceivers would also permit more ports to be added to enterprise servers, fiber-optic switches, and communications equipment without increasing the size and cost of these devices [2, 3]. Recently, a new class of small form factor (SFF) fiber-optic connectors have been introduced with the goal of reducing the size of a fiber-optic connector to one-half that of an SC Duplex connector while maintaining or reducing the cost [4], namely, the LC [5], MT-RJ [6], SC-DC [7], and VF-45 [8].

Various types of next-generation SFF optical interfaces have been proposed to the Electronics Industry Association/Telecommunications Industry Association (EIA/TIA), for inclusion in developing standards such as the Commercial Cabling Standard TIA-568-B. These connectors are described by a reference document called a Fiber-Optic Connector Intermatability Standard (FOCIS), which defines the connector geometry so that the same connector build by different manufacturers will be mechanically compatible. The EIA/TIA also requires connectors to

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

Table 3.1
EIA/TIA Proposed SFF Connector Standards.

Connector Type	FOCIS Document Number	FOCIS Author*
MT-RJ	12	Amp
SC-DC	11	Siecor
LC	10	Lucent
SG (VF-45)	7	3M
Fiber Jack	6	Panduit

Note: the authoring company listed does not necessarily support or manufacture only one connector type, nor does this table include all of the supporters or manufacturers of each connector type.

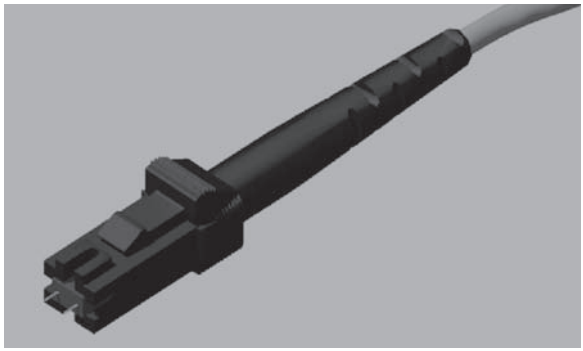


Figure 3.1 Standard MT-RJ male connector.

meet some minimal performance levels, independent of connector design; test methodologies are defined by the EIA/TIA Fiber-Optic Test Procedures (FOTPs). Other industry specifications are also relevant to these connectors; for example, Bellcore spec. GR-326-CORE defines fiber protrusion from a ferrule to ensure physical contact and to prevent backreflections in the connector. The relevant FOCIS documents defined for connectors that have currently been proposed to the TIA are given in Table 3.1.

3.2. MT-RJ CONNECTOR

The MT-RJ connector utilizes the same rectangular plastic ferrule technology as the Multi-fiber Termination Push-on (MTP) array style connector first developed by NTT Corporation, with a single ferrule body housing two fibers at a 750-μm pitch as in Fig. 3.1. These ferrules are available in both single-mode and multimode tolerances, with the lower cost multimode version typically comprised

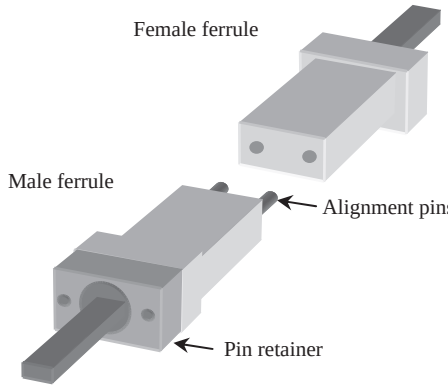


Figure 3.2 MT-RJ Ferrule alignment method.

of a glass-filled thermoplastic and the critically tolerance singlemode version comprised of a glass-filled thermoset material. Unlike the thermoplastic multimode ferrules, which can be manufactured using the standard injection mold process, the thermoset single-mode ferrules must be transfer molded, which is generally a slower but more accurate process.

By design, the alignment of two MT-RJ ferrules is achieved by mating a pair of metal guide pins with a corresponding pair of holes in the receptacle (Fig. 3.2). This feature makes the MT-RJ the only small form factor connector with a distinct male and female connector. As a general rule, wall outlets, transceivers, and internal patch panel connectors will retain the guide pins, thus making their gender male, and the interconnecting jumpers will have no pins (female). In the event that the two jumper assemblies require mating mid-span, a special cable assembly with one male end and one female end must be used. However, some unique designs do allow for the insertion and extraction of guide pins in the field, affording the user the ability to change the connector's gender as required.

Latching of the MT-RJ connector is modeled after the copper RJ-45 connector, whereby a single latch arm positioned at the top of the connector housing is positively latched into the coupler or transceiver window. Although this latch design is similar in all the MT-RJ connector designs, individual latch pull strengths may vary depending on the connector material, arm deflection, and the relief angles built into the mating receptacles. For this reason it is recommended that connector pull strengths be evaluated as a complete interface; depending on the specific manufacturer's connector, coupler, or transceiver design, the coupling performances may vary.

MT-RJ connectors are typically assembled on 2.8-mm round jacketed cable housing two optical fibers in one of three internal configurations. The first construction style consists of the two optical fibers encapsulated within a ribbon at

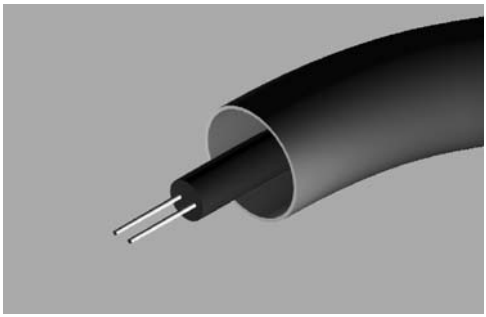


Figure 3.3 900-micron buffer with 250-micron fibers.

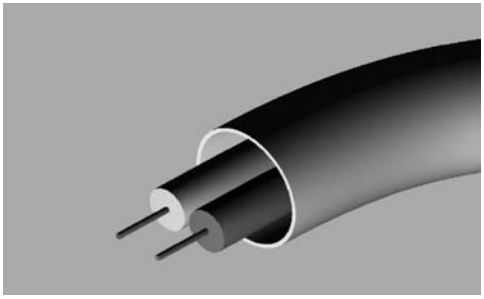


Figure 3.4 Dual 900-micron buffered fibers.

a 750-μm pitch (Fig. 3.3). This approach is unique to the MT-RJ connector and is designed specifically to match the fibers spacing to the pitch of the ferrule for ease of fiber insertion. Although this construction style may be ideal for a MT-RJ termination, it does present some difficulty when manufacturing a hybrid assembly and availability may also be an issue based on its uniqueness. A second design, which is more universal, utilizes a single 900-μm buffer to house two 250-μm fibers (Fig. 3.4). This construction is more conducive to hybrid cable manufacturing, but the fibers will naturally maintain a 250-μm pitch, thus making fiber insertion rather difficult.

The third design is considered a standard construction and is used across the industry. In this configuration each individual fiber is buffered with a PVC coating. The coating thickness is typically 900 μm, but as in the previous case this does cause a mismatch of the fiber to ferrule pitch. To compensate for this, some connector designs incorporate a fiber transition boot, which gradually reduces the fiber pitch to 750 μm, while others simply use a nonstandard buffer coating of 750 μm.

In general, the assembly and polish of the MT-RJ factory style connector is considerably more difficult than the other small form factor connectors. Typical MT-RJ designs have a minimum of eight individual components that must be assembled after the ferrule has been polished, allowing for a number of handling concerns. As with the case of most MT-style ferrules, the perpendicularity or flatness of the ferrule end face with reference to the ferrule's inner shoulder is critical, and this cannot be accomplished if the connector is preassembled. Another unique requirement of the MT-RJ polish involves fiber protrusion. The ferrule end face is considered to be flat, but depending on the polishing equipment, fixtures, and even contamination, some angularity may occur. Therefore, it is recommended that the fibers themselves protrude 1.2 to 3.0 μm from the ferrule surface in order to guarantee fiber-to-fiber contact.

3.3. SC-DC

The SC-DC[1] (dual contact) or SC-QC (Quattro contact), developed by Siecor and illustrated in Fig. 3.5, has a connector body design resembling a SC simplex connector with a round thermoset-molded ferrule capable of handling either two or four optical fibers, as in Fig. 3.5. The connector is designed to support both single-mode and multimode cabling applications. However, the SC-DC is one of the few connectors that is used exclusively for cable interconnecting and therefore has no transceiver support.

As in the case of the MT-RJ connector, the SC-DC fibers are on a 750-μm pitch, while the SC-QC fibers are on a 250-μm pitch. The same ferrule and

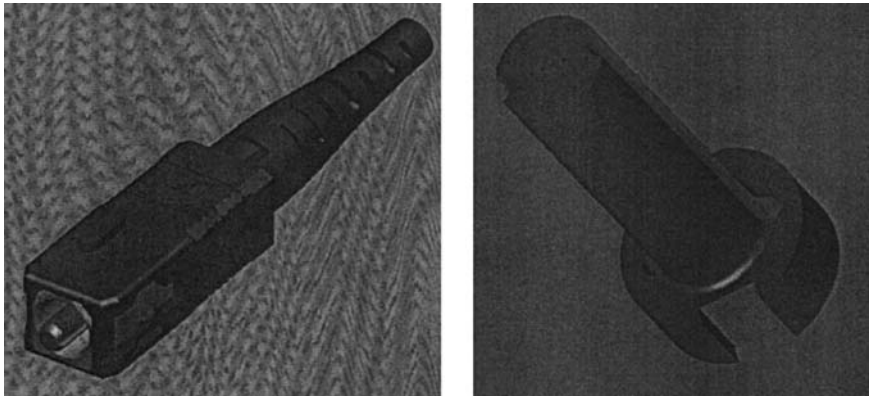


Figure 3.5 Siecor SC-DC/SC-QC Connector and ferrule assembly.

[1]SC-DC and SC-QC are trademarks of Siecor Corporation.

housings are used in both connectors, so the only feature that distinguishes between the two is simply the number of fibers used. The four ferrule holes of approximately 126-μm diameter are placed on 250-μm centers along the ferrule centerline to form a SC-QC connector. By using only the two outermost ferrule holes and leaving the inner two empty, the SC-DC is created.

Although the ferrule composition is very similar to that used across the MT technologies, the geometry and alignment methods are very different. The SC-DC ferrule has a standard round shape with a 2.5-mm diameter, but unlike its ceramic counterparts, there are two semicircular groves with a 350-μm radius positioned along each side of the ferrule at a 180° interval. This feature provides the ferrule-to-ferrule alignment when mated with the corresponding ribs of the coupler. The design of this feature within the couplers is different for a single-mode and multimode connection. The multimode coupler uses a one-piece, all composite alignment insert with molded ribs, while the single-mode coupler incorporates two precision alignment pins captured inside the sleeve. By design, the ribs of the coupler and grooves in the ferrule are on the same 2.6-mm pitch as the pins of a male MT-RJ connector to allow for possible hybrid mating of the two connector types.

The latching mechanism of the SC-DC connector uses the same push-pull style used on the industry standard SC simplex connector, and the outer housing dimensions are identical. Because the two connectors appear the same physically but are not functionally interchangeable, the housing alignment key of the SC-DC is offset to help distinguish between the two connector types and prevent any confusion in the field. By following the same basic footprint as the SC connector, this new SFF optical connector already has a familiar look and feel with proven plug reliability.

SC-DC connectors and cable assemblies are manufactured and sold solely by Corning; therefore, the variability of design and cable material is not an issue as is the case with other SFF connectors. The connectors are typically assembled onto 2.8-mm round jacketed fiber incorporating a single-ribbon fiber populated with either two or four fibers, which assists in the alignment of the fiber to the ferrule holes. The internal ferrule geometry can accommodate the standard single-fiber cables, and this may be an option for the fabrication of hybrid cable assembles.

As previously stated, the size and shape of the SC-DC ferrule is similar to the standard ceramic and therefore can be polished in much the same manner and still produce endface geometries akin to the MT-RJ. A flat endface polish is the desired result, much as it is with the MT-style ferrules. However, the perpendicularity is referenced to the sides of the ferrule rather than to an internal feature. Therefore, the SC-DC connector can be preassembled and polished as a complete connector. The ability to preassemble a connector significantly reduces the complexity of manufacturing and typically results in a low rework rate.

The SC-DC connector is available in a field-installable version, the SC-DC UniCam, which utilizes pre-polished fiber stubs much like other SFF solutions. Alignment of the fiber stubs to the in-field cleaved fiber is achieved through a gel-filled mechanical splice element. The splice element is opened and closed with a mechanical cam and is retained within a connector housing. Termination of the standard SC-DC connector can also be accomplished in the field by using conventional equipment and methods much like the field termination of a SC-style connector.

3.4. VF-45

The VF-45[2] connector, developed by 3M and shown in Fig. 3.6, is perhaps the most innovative SFF connector design in that it eliminates the need for precision ferrules and sleeves altogether. The overall look and feel of the "plug-to socket" design closely resembles the standard telephony RJ-45 "connector-to-jack" system whereby the cable assembly mates directly to a terminated socket, therefore reducing the need for couplers. Although this concept has been around for years in the cooper industry, the creation of a bare fiber-optical interface, using alignment grooves and no index matching gels, requires some revolutionary techniques.

The VF-45 connector incorporates two 125-μm optical fibers, suspended in free space. on a 4.5-mm pitch and protected by a RJ-45 style housing with a retractable front door designed to protect the fibers. The connector design supports both single-mode and multimode tolerances by relying on the inherent precision

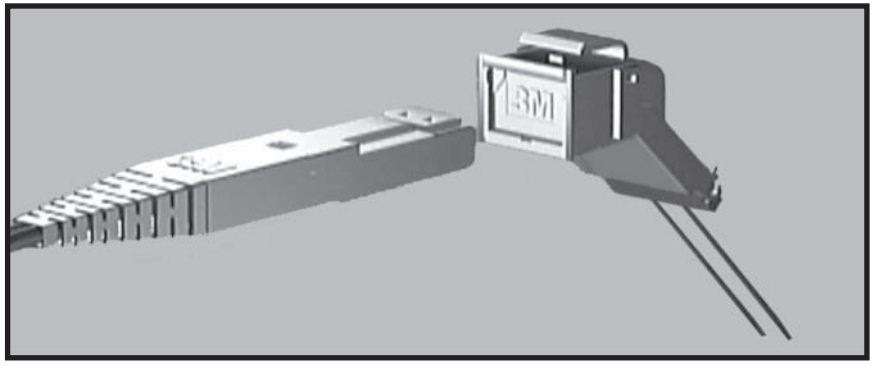


Figure 3.6 3M VF-45 "Connector-to jack" system.

[2]VF-45 is a trademark of 3M Corporation.

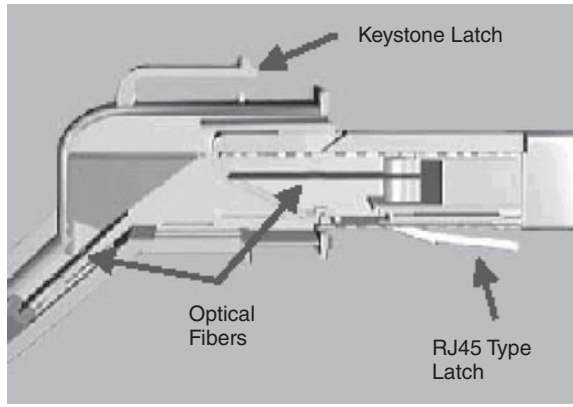


Figure 3.7 Insertion of VF-45 plug.

of the optical fibers within the two injected molded v-grooves of either the transceiver or a VF-45 socket. The design of the interconnect allows the natural spring forces of the optical fibers to align the fibers within the v-grooves as well as ensuring physical fiber-to-fiber contact.

Because of the uniqueness of this interconnect, the geometry of the endface polish of both the plug and receptacle fibers have been modified to provide optimum performance. The VF-45 optical connection relies on the spring force created by the bowing of the optical fibers to provide a physical contact force of approximately 0.1 N. This force, coupled with an 8-degree angle polished endfaces, produces the optimum connection and return loss results. The tips of the plug fibers are also beveled at 35 degrees, allowing a 90-μm contact area and providing a relief for the fiber to slide into the v-grooves with no damage to the core region (see Figs. 3.7 and 3.8). This chamfer is not required on the receptacle fibers since they remain stationary while the plug fibers may have to endure multiple insertions.

As previously mentioned, the contact force created at the optical interface directly influences the optical performance of the plug-socket connection. This downward compressive force is generated when the two fibers of the plug engage with the resident fibers of the socket and cause a slight "bow" in the plug fibers. Because of the constant stress on these fibers, long-term reliability on standard optical fibers became a concern, and therefore a specialized high-strength optical fiber was developed for this application, called GGP (glass-glass-polymer) fiber. GGP fiber consists of 100-μm glass fiber with a polymeric coating applied to bring the outer diameter to 125 μm. By reducing the outer diameter of the glass, the tensile stress on the fiber is minimized and the additional coating also provides protection against abrasions from the v-grooves and reduces the chance of damage

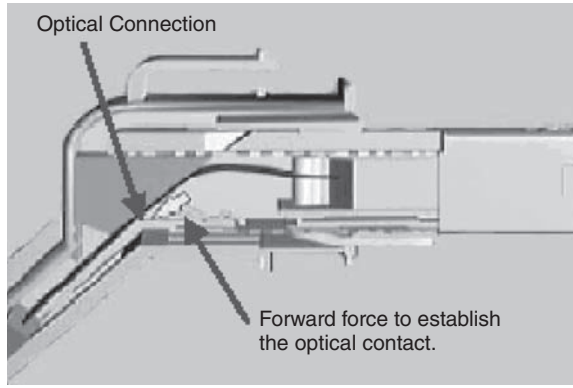


Figure 3.8 VF-45 optical connection.

to the glass during the mechanical stripping process used in the termination process.

The factory termination of the VF-45 jumper plugs is considerably different from the conventional ferrule-based connectors. The process of threading a 125-μm fiber into a precision ferrule hole filled with epoxy is now eliminated and replaced with a mechanical fiber holder that grips the fibers in place. The fibers are then cleaved and polished to the endface geometry previously described, and the cable strain relief is slid into place. The fibers and holder are then placed in a protective shroud, and the front door cover is installed. The relative simplicity of this manufacturing process makes the VF-45 connector one of the best candidates for a fully automated production line.

The socket of the VF-45 was specifically designed for termination in the field with minimal effort and training. After preparing the fibers for termination by removing outer buffer material, they are inserted into a mechanical fiber holder that retains the fiber by gripping them inside a deformable aluminum crimp. The fibers are then cleaved and hand polished to an 8-degree angle, with a slight radius generated by the durameter of the polishing pad. The fiber holder with the polished fibers is guided in to the socket v-grooves, and the housing plate is snapped into place (see Fig. 3.9). Although this method of field termination does vary from the other pre-polished SFF connectors, the total termination time and the complexity of the process are very similar.

3.5. LC CONNECTOR

The LC connector developed by Lucent Technologies and shown in Fig. 3.10 is a more evolutionary approach to achieving the goals of a SFF connector. The

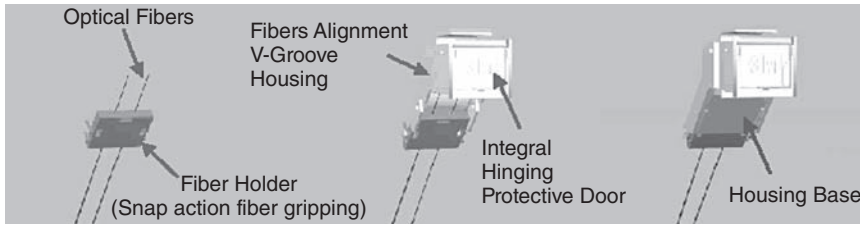


Figure 3.9 Field termination of the 3M VF-45 socket.

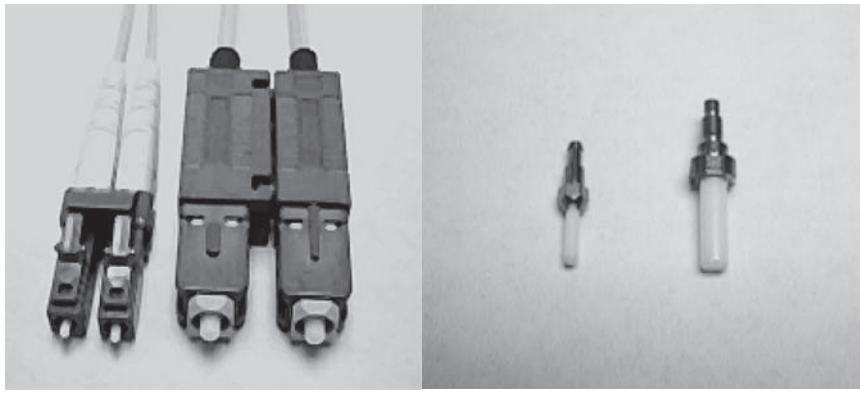


Figure 3.10 LC duplex connector and ferrule compared with a SC duplex connector and ferrule.

LC connector utilizes the traditional components of a SC duplex connector having independent ceramic ferrules and housings, with the overall size scaled down by one-half (Fig. 3.10). The LC family of connectors includes a stand-alone simplex design; a "behind the wall" (BTW) connector and the duplex connector available in both single-mode and multimode tolerances are all designed using the RJ-style latch.

The outward appearance and physical size of the LC connector varies slightly depending on the application and vendor preference. Although all the connectors in the LC family have similar latch styles modeled after the copper RJ latch, the simplex version of the connector has a slightly longer body than either the duplex or BTW version, and the latch has an additional latch actuator arm that is designed to assist in plugging as well to prevent snagging in the field. The BTW connector is the smallest of the LC family and is designed as a field- or board-mountable connector using 900-μm buffered fiber and in some cases has a slightly extended latch for extraction purposes. The duplex version of this connector has a modified

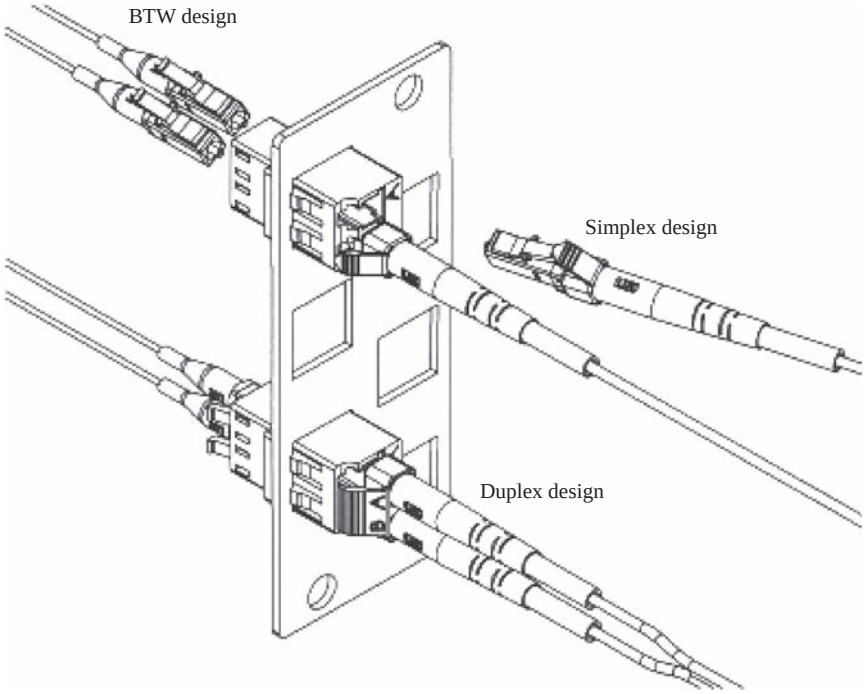


Figure 3.11 The family of LC style connectors.

body to accept the duplexing clip that joins the two connector bodies together and actuates the two latches as one (see Fig. 3.11). Finally, even the duplex clip itself has variations depending on the vendor. In some cases the duplex clip is a solid one-piece design and must be placed on the cable prior to connectorization, while other designs have slots built into each side to allow the clip to be installed after connectorization. In conclusion, all LC connectors are not created equal, and depending on style and manufacturer's preference, there may be attributes that make one connector more suitable for a specific application then another.

The LC duplex connector incorporates two round ceramic ferrules with outer diameters of 1.25 mm and a duplex pitch of 6.25 mm. These ferrules are aligned through the traditional couplers and bores using precision ceramic split or solid sleeves. In an attempt to improve the optical performance to better than 0.10 db at these interfaces, most of the ferrule and backbone assemblies are designed to allow the cable manufacturer to tune them. Tuning of the LC connector simply consists of rotating the ferrule to one of four available positions dictated by the backbone design. The concept is basically to align the concentricity offset of each

ferrule to a single quadrant at 12:00; in effect, if all the cores are slightly offset in the same direction, the probability of a core-to-core alignment is increased and optimum performance can be achieved. Although this concept has its merits, it is yet another costly step in the manufacturing process, and in the case where a tuned connector is mated with an untuned connector, the increase in performance may not be realized.

Typically, the LC duplex connectors are terminated onto a new reduced-size zipcord referred to as mini-zip. However, as the product matures and the applications expand, it may be found on a number of different cordages. The mini-zip cord is one of the smallest in the industry with an outer diameter of 1.6 mm compared with the standard zipcord for an SC style product of 3.0 mm. Although this cable has passed industry standard testing, the cable manufacturers have raised some issues concerning the ability of the 900-μm fibers to move freely inside a 1.6-mm jacket and others involving the overall crimped pull strengths. For these reasons, some end users and cable manufactures are opting for a larger 2.0-mm, 2.4-mm, or even the standard 3.0-mm zipcord. In applications where the fiber is either protected within a wall outlet or cabinet, the BTW connector is used and terminated directly onto the 900-μm buffers with no jacket protection.

The factory termination of the LC cable assemblies is very similar to other ceramic-based ferrules using the standard pot and polish processes with a few minor differences. The one-piece design of the connector minimizes production handling and helps to increase process yields when compared with other SFF and standard connector types. Because of the smaller diameter ferrule, the polishing times for an LC ferrule may be slightly lower than the standard 2.5-mm connectors, but the real production advantage is realized in the increased number of connectors that can be polished at one time in a mass polisher. For the reasons mentioned above and because the process is familiar to most manufacturers, the LC connector may be considered one of the easiest SFF connectors to factory terminate.

Field termination of the LC connector has typically been accomplished through the standard pot and polish techniques using the BTW connector. However, a pre-polished, crimp and cleave connector is also available. The LCQuick Light field-mountable BTW style connector made by Lucent Technologies is a one-piece design with a factory polished ferrule and an internal cleaved fiber stub. Unlike other pre-polished SFF connectors previously discussed, the LCQuick light secures the inserted field cleaved fiber to a factory polished stub by crimping or collapsing the metallic entry tube onto the buffered portion. This is accomplished by using a special crimp tool that is designed not to damage the fibers. However, this means the installer has but one chance for a good connection. The LCQuick light is designed specifically for use in protected environments such as cabinets and wall outlets and has no provision for outer jacket or Kevlar protection.

3.6. OTHER TYPES OF SFF CONNECTORS

Following a renewed interest in fiber-optic cabling and transceiver footprint reduction, many types of small form factor optical interfaces have been proposed. Some of these are no longer widely used; for example, the mini-MT connector was originally proposed as a 2-fiber version of the MTP connector, but was largely supplanted by the MT-RJ design. In this chapter, we have concentrated on optical interfaces for next-generation transceivers and cable plant infrastructures, and we have limited our treatment so as not to include the MTP/MPO connectors (see Chapter 24), the SMC parallel optical connector (see Chapter 11), or other emerging multifiber interfaces for Infiniband. However, some other SFF interfaces are currently in use besides the four major types discussed earlier; we will briefly describe them here.

3.6.1. Fiber-Jack

The Fiber-Jack (FJ) is the nontrademarked name for the Opti-Jack connector developed by Panduit Corporation. As shown in Fig. 3.12, this connector incorporates two industry standard SC duplex ceramic ferrules, each 2.5 mm in diameter. However, the spacing between ferrules has been reduced from 1.27 mm (0.5 in.) as in a standard SC Duplex connector to only 0.63 mm (0.25 in.). The ferrules are independently spring loaded and are aligned in a receptacle by standard split-sleeve mechanical techniques. While this simplifies the connector design, it also means that the connector must incur the full cost of two ceramic ferrules; how this will eventually compare with other SFF ferrule costs has yet to be determined. The connector latch is modeled after the industry standard RJ-45 wall jack and has found many initial applications in building wiring to wall outlets. The connector is available in both single-mode and multimode versions, which preserve the TIA industry standard color coding on the plug body and the termination cap on the jack.

Figure 3.12 The Fiber-Jack Connector.

The FJ connector is also available with color coding to identify different networks, applications, areas of the building, or portions of the cable infrastructure, to facilitate network adminstration. Following the category 5 wiring conventions, the FJ housing and plugs may be color coded in black, blue, gray, orange, red, beige, or white. The FJ was also among the first SFF connectors to be specified by a TIA FOCIS document for use with plastic optical fiber. Because it employs standard ceramic ferrules, the FJ can accomodate plastic fiber with an outer diameter up to 1 mm. The FJ supports standard duplex jumper cables, couplers, and adapters, although FJ transceivers are not widely available.

3.6.2. MU

Another type of SFF connector, developed by NTT to serve as both a standard fiber-optic patch cable and an optical backplane interface, is the multitermination unibody or MU connector. This connector is also available from various sources under slightly different names; for example, the version manufactured by Sanwa Corporation is called the SMU. As shown in Fig. 3.13, the basic MU connector is a simplex design measuring 6.6 mm wide and 4.4 mm high, with a center-to-center spacing of 4.5 mm in duplex or multifiber applications. It was standardized by IEC 61754-6, "Interface standards type MU connector family," in 1997; other standards bodies, including JIS and IEEE 1355 Heterogeneous Interconnect (a future bus architecture), have endorsed the MU as well. A backplane version of the MU is available, which measures 13 mm wide and 42 mm high. Its small size is achieved by using a ceramic ferrule 1.25 mm in diameter, roughly half the size of a standard SC connector ferrule. Consequently, a different type of physical contact polishing was developed to accomodate the smaller ferrules.

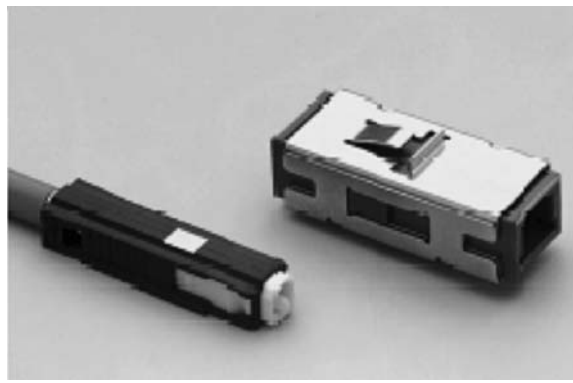


Figure 3.13 The MU connector.

Smaller diameter zirconia split-sleeves have also been developed to support duplex couplers, adapters, and similar jumper cable applications. A self-retention mechanism is employed, similar in design to a miniature push-pull SC latch. Indeed, the MU is sometimes referred to as a mini-SC connector because of the similarities in look and feel. This is consistent with the intended applications, as it permits blind mating of the connector in a printed circuit board backplane more readily than an RJ-45 style latch.

The MU has found some applications as a front panel patch cable on optical switches or multiplexers that have a large amount of fiber connections; these interfaces are expected to be incorporated into optical backplanes on next-generation equipment. The plug and jack structure of MU connectors can be easily cascaded to produce multifiber parallel interface designs. It has been suggested that future designs using the MU could accommodate hundreds or even thousands of fibers when used in this configuration. Transceivers for the MU interface are not yet widely available, although some development work is under way in this area.

3.7. SFF, SFP, AND SFP + TRANSCEIVERS

The industry trend has been to adopt the MT-RJ interface for low data rate (below 1 Gbit/s), multimode applications, and the LC for high data rate (above 1 Gbit/s) applications, both single-mode and multimode (mainly using SX transceivers). Transceiver development has been facilitated by ad hoc industry standards or multisource agreements (MSAs), which govern transceiver package dimensions, electrical interfaces and host board layouts, card bezel design, mechanical specifications (including insertion, extraction, and retention forces), and transceiver labeling. The original MSA for SFF transceivers was supported by 15 companies, including Agilent, IBM, Lucent, Siemens/Infineon, Amp/Tyco, and others. It defined a pin through hole device with two rows of five pins each.

Recently, a second MSA has been approved by the member companies and defines a pluggable transceiver that mates with a surface mountable card receptacle. These small form factor pluggable (SFP) transceivers make it possible to change the optical interface at the last step of card manufacturing, or even in the field, to accommodate different connector interfaces or a mix of SX and LX transceivers. This should make it easier to adjust optical interface characteristics on future system designs, in much the same way that the GBIC transceiver did for the SC Duplex interface (in fact, the SFP is sometimes known as a “mini-GBIC”). The SFP has 20 signal connections and provides three additional functions in addition to the original 10 SFF signal pins. These new functions include module definition pins that specify a serial ID indicating the type of transceiver function (such as LX vs. SX transmitters),

a data rate select function (such as 1 Gbit/s vs. 2 Gbit/s), and a transmitter fault signal.

A metal receptacle, sometimes called a cage or a garage, is surface mounted to the printed circuit board to accept the pluggable transceivers. In addition to providing easy replacement and reconfiguration of the transceiver interface, this offers several other advantages. The transceiver is often the only pin through hole component on a modern card design; the SFP cage allows the elimination of extra manufacturing processing steps and potentially reduces cost. By removing the optical components from the soldering process, the SFP should provide improved reliability of the optics and permits the use of higher soldering temperatures. (This may be important as environmental regulations require future lead-free solders with higher process temperatures.)

The industry has recently developed enhancements to the SFP MSA, known as SFP Plus (SFP +), which is intended to achieve higher data rates, lower cost, and improved thermal performance. As of this writing, this specification has not been finalized for public release, although many companies are now designing compliant transceivers. Although SFP+ is applicable to the same application set as SFP, and includes both copper and optical interfaces, it is particularly intended to support high data rate links such as 8.5 and 10.52 Gbit/s Fibre Channel, 10 Gbit/s Ethernet (10.31 Gbit/s links, or 11.1 Gbit/s links using forward error correction, including 10GBase SR, LR, and LRM), SONET OC-192 (9.95 Gbit/s), and G.709 "OTU-2" (10.7 Gbit/s). The specification is similar to SFP, with a common form factor and optical interface. There is a new electrical interface specification, called SFI, designed to handle higher data rate performance; this is defined in SFP specification 8431. In particular, the jitter budgets are planned to be somewhat tighter for SFI than for a standard SFP interface (though not as restrictive as the XFI interface). We also note that SFP+ has not defined a data rate selection pin, meaning that SFP+ transceivers may be compatible with lower data rates but not necessarily compliant with the older specifications.

A revised SFP+ mechanical specification (sometimes known as "improved pluggable form factor"), including thermal and electromagnetic compatibility, is also available through SFP specification 8432. In particular, SFP+ defines two classes of maximum power dissipation: class I (up to 0.8 W) and class II (up to 1.5 W). The class II transceivers are intended for DWDM and telecom applications (single-height cage with cooled optics). The SFP+ transceivers are backward compatible with most (but not necessarily all) SFP cages implemented according to the SFF-8074i specification. However, in this case the full benefit of SFP+ improvements in electromagnetic susceptibility and other parameters may not be achievable.

At this writing, some important practical questions remain, such as whether the 8G and 10G implementations can be realized with a common transceiver. Although they will use the same optics, 8G Fibre Channel needs to be backward

compatible with at least 2 previous generations of the Fibre Channel standard (4G and 2G) using 8B/10B encoding. (As noted in Chapter 20, 10G Fibre Channel uses a different encoding scheme and is therefore not backward compatible with lower data rates.) By contrast, to accommodate the full range of Ethernet standards options including FEC, the 10G transceiver must operate over data rates ranging from 10.3 to 11.1 Gbit/s using 64B/66B encoding. The use of a single electrical receiver design (linear vs limiting) for 10G remains a design issue (see Chapter 20 and case study), as well as the question of whether clock and data recovery should be more closely integrated into the transceiver.

3.7.1. QSFP

Recent interest has been shown in further increasing the transceiver port density for optical datacom applications, particularly blade servers, as well as finding a cost effective way to increase aggregate data rates beyond 10 Gbit/s. One approach is the quad small form factor pluggable (QSFP) multisource agreement, originally announced in March 2006 and finalized in December 2006 following a period of public comment on the proposed specifications. The QSFP MSA (www.qsfpmsa.org) is currently endorsed by over 20 companies. As illustrated in Fig. 3.14, QSFP defines an integrated, hot pluggable, four-channel optical transceiver, designed to replace four standard SFP modules in a space only 30% larger than a single SFP. The resulting port density is three times higher than conventional SFP designs; various stacked and ganged configurations are possible to achieve increased port density, and presumably lower cost per port (a minimum of 21 mm center-to-center spacing is allowed for adjacent QSFP transceivers). The transceiver is a so-called z-axis pluggable module, meaning that the 38 contact electrical connector can be inserted parallel to the host circuit board without requiring additional operations such as screwing the transceiver package to the host card.

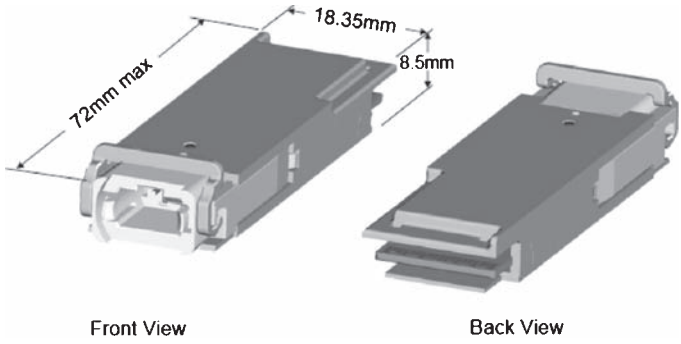


Figure 3.14 QSFP transceiver.

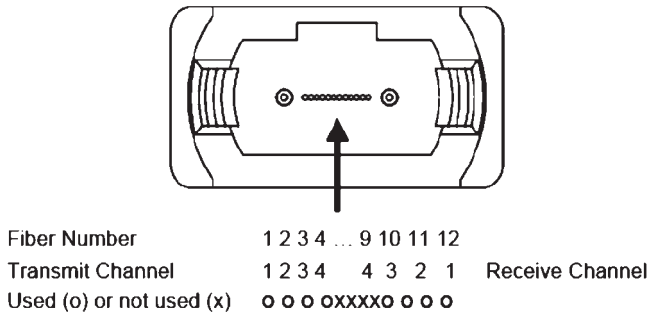


Figure 3.15 Definition of MPO connector interface for QSFP (viewed looking back into the transceiver port, keys on top).

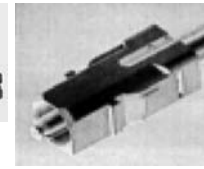
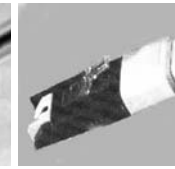
QSFP accommodates a standard MPO connector, though as shown in Fig. 3.15 only 8 of the 12 available fibers are used to carry signals. Exposed portions of the optical connector or card bezel are color coded following common industry practice (beige for 850 nm/multimode, blue for 1300 nm/single-mode, white for 1550 nm/single-mode). The MSA defines a mechanical form factor with latching mechanism similar to that used by XFI, host board electrical receptacle, and a cage to house the transceiver when it is plugged onto a host card. Digital diagnostics are provided to monitor link performance; the diagnostic interface includes the ability to set link distance parameters that identify the capabilities of the transceiver, including an option for copper QSFP implementations. QSFP uses a single 3.3 V nominal power supply, with a maximum power dissipation under 3.5 W. Various types of vendor-specific heat sinks can be attached or clipped onto the transceiver. Data rate options ranging from 100 Mbit/s to 10 Gbit/s per channel are defined (with an available rate select pin on the electrical interface) to support protocols, including Ethernet, Fibre Channel, InfiniBand, and SONET/SDH. In particular, with a potential data rate of 10 Gbit/s/channel, this transceiver may provide a cost-effective implementation of 40 Gbit/s links. The inherent 4 + 4 channel architecture of QSFP lends itself to increase distances supported by multilane serial I/O electrical interconnects such as PCI Express (PCIe) and InfiniBand [9].

3.8. COMPARISON OF SFF FORM FACTORS

Table 3.2 presents a comparison of the different features of the four major SFF connectors. A brief description of each connector and its alignment method are given, followed by a discussion of the distinguishing characteristics and their impact on the connector and transceiver.

Several different design approaches can be used to reduce the dimensions of a fiber-optic connector. One approach is to use a single ferrule with multiple

Table 3.2
Comparison of SFF Connector Features.

	LC	MT-RJ	SC-DC[vii]	VF-45[viii]
				
Fiber spacing	6.25 mm	0.75 mm	0.75 mm	4.5 mm
No. of ferrules	2	1	1	0
Ferrule material	Ceramic	Plastic	Plastic	None
Alignment	Bore & Ferrule	Pin and Ferrule	Rail and Ferrule	V-groove
Ferrule Size	ϕ 1.25 mm	2.5 mm × 4.4 mm	ϕ 2.5 mm	None
Trx opening: (width × height × length)	11.1 mm × 5.7 mm × 14.6 mm	7.2 mm × 5.7 mm × 14 mm	11 mm × 7.5 mm × 12.7 mm	12.1 mm × 8 mm × 21 mm
Fiber cable	duplex	duplex or ribbon	duplex or ribbon	GGP polymer coated
Field term: Plug	pot & polish	pre-polished stub	pre-polished stub	Not Avail
Field term: Socket	plug + coupler	plug + coupler & socket	plug + coupler	cleave & polish socket
Latch	RJ—top 2 latch coupled	RJ—top latch	SC push pull	RJ—top latch

fibers; this is the concept behind the SC-DC and MT-RJ connectors. The SC-DC (dual connect) and SC-QC (quad connect) use a standard SC connector body and latching mechanism with an offset key, but a new round plastic ferrule design that incorporates either two fibers (750-mm pitch) or four fibers (250-mm pitch) in a linear array. Alignment is provided by semicircular grooves in the sides of the SC-DC ferrule, which mate with corresponding ribs in the receptacle. This connector has been used by IBM Global Services as part of the Fiber Quick Connect system; it is currently limited to applications in patch panels and the cable infrastructure. The most radical, and innovative, approach for a smaller connector is to eliminate ferrules altogether; this is the case for the VF-45 connector. In this connector, a pair of optical fibers are aligned using injection-molded thermoplastic V-grooves; the fibers are cantilevered in free space on a 4.5-mm pitch and are protected by the connector outer body. When plugged into a receptacle, the fibers bend slightly in order to achieve physical contact; better performance is achieved when using optical fibers that have a special strength coating in addition to the outer jacket.

The MT-RJ connector uses the same rectangular plastic ferrule concept as the multifiber MTP connector, with two fibers on a 750-mm pitch and a latching mechanism based on the RJ-45 connector. Alignment in this case is provided by a pair of metal guide pins in the connector, which mate with a corresponding pair of holes in the receptacle. This feature makes the MT-RJ the only small form factor connector with distinct male and female connector ends. A more evolutionary approach to designing SFF connectors involves simply shrinking the standard SC Duplex connector, maintaining a single fiber in each of the ceramic ferrules, and using conventional alignment techniques applied to the ferrules. The LC connector uses this approach and shrinks the ferrules to 1.25 mm in diameter with a fiber pitch of 6.25 mm (duplex). LC is the only small form factor connector which can be either simplex or duplex.

Although many of these designs exhibit comparable performance, some application-specific differences remain, such as axial pull requirements. Some SFF connectors, such as the SC-DC and VF-45 connectors, are designed to disengage from their receptacles under an applied pull force above 45 N; this is to prevent damage to the connector and induce an obvious optical failure in the link. (It is problematic to diagnose a link failure if the connector remains engaged but exhibits high optical losses under stress.) Another design approach used by LC and MT-RJ is to retain the connector in the receptacle under much higher force without excessive optical loss. The magnitude of the pull force is a requirement that will vary depending on the application. However, it is essential for all the applications that the cable unplug or mechanically fail before the loss increases.

REFERENCES

1. ANSI Fibre Channel Standard—Physical and Signalling Interface (FC-PH) X3.230 rev 4.3. 1994.
2. Trewhella, J., C. DeCusatis, and J. Fox. 2000, July. Performance comparison of small form factor fiber optic connectors, *IEEE Transactions on Advanced Packaging* 23, no. 2:188–196.
3. DeCusatis, C., J. Trewhella, and J. Fox. 1999, December. Small form factor optical fiber connectors: performance comparison. *Optics & Photonics News*, p 34.
4. Schwantes, C. 1998, October. Small-form factors herald the next generation of optical components. *Lightwave* pp. 65–68.
5. Shahid, M. A., et al. 1999, June. Small and efficient connector system. *Proc. 49th Electronic Components and Technology Conf.*, San Diego, Calif., and TIA FOICS 10.
6. Tamaki, Y., et al. 1999, June. Compact and durable MT-RJ connector. *Proc. 49th Electronic Components and Technology Conf.*, San Diego, and TIA FOICS 12.
7. Wagner, K. 1998, December. SC-DC/SC-QC Connector, *Optical Engineering* 37, 12:3129 and TIA FOICS 11.
8. Selli, R., et al. 1998, December. A novel V-groove based interconnect technology. *Optical Engineering* 37, 12:3134 and TIA FOICS 7.
9. Kipp, S., ed. 2006, December 4. Quad small form factor pluggable (QSFP) transceiver specification, rel. 3.0.

4

Specialty Fiber-Optic Cables

Casimer DeCusatis
IBM Corporation, Poughkeepsie, N.Y.

John Fox
ComputerCrafts Inc., Hawthorne, N.J.

4.1. INTRODUCTION: CONVENTIONAL OPTICAL FIBER

In this chapter, we will describe different types of fiber-optic cable that have been developed for a wide range of applications, including enhanced distance, optical amplification and attenuation, dispersion and polarization management, and other areas. The fabrication of conventional low-loss silica fiber-optic cables [1] involves precision control of the glass composition to within several parts per million (ppm). The desired refractive index profile is first fabricated by selectively doping a large glass preform, or boule, typically several centimeters in diameter and about a meter long, which maintains the relative dimensions and doping profiles for the core and cladding. The boule is later heated in an electric resistance furnace until it reaches its melting point over the entire cross section; it can take up to an hour to establish uniform heating of the preform. Thin glass fibers are then drawn upward from the preform in a drawing tower; the fiber cools and solidifies very quickly, within a few centimeters of the furnace. The pulling force controls the rate of fiber production, and hence the fiber diameter, which is monitored by a laser interferometer. Bare glass fiber is then drawn through a vat of polymer and receives a protective coating extruded over it to a diameter of about 250 microns; this must be done as soon as possible after drawing to avoid water contamination in the fiber. Finally, the fibers are spooled evenly onto a mandrel about 20 cm in diameter, to avoid microbending. If the preform is uniformly heated (and therefore has a uniform viscosity), then the cross section and index profile of the drawn fiber will be exactly the same as in the preform. In this manner, fibers with very complex refractive index profiles can be produced. The

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

primary technology used in fiber preform manufacturing is chemical vapor deposition (CVD), in which submicron silica particles are produced through one or both of the following chemical reactions, carried out at temperatures of around 1800 to 2000°C:

$$\begin{aligned} SiCl_4 + O_2 &\rightarrow SiO_2 + 2Cl_2 \\ SiCl_4 + 2H_2O &\rightarrow SiO_2 + HCL \end{aligned} \tag{4.1}$$

This deposition produces a high-purity silica soot that is then sintered to form optical quality glass. Two basic manufacturing techniques are commonly used. In the so-called inside process, a rotating silica substrate tube is subjected to an internal flow of reactive gasses. Two variations on this approach are modified chemical vapor deposition (MCVD) and plasma-assisted chemical vapor deposition (PCVD). In both cases, layers of material are successively deposited, controlling the composition at each step, in order to reach the desired refractive index. MCVD (which accounts for a large portion of the fiber produced today, especially in America and Europe) accomplishes this deposition by application of a heat source, such as a torch, over a small area on the outside of the silica tube. Submicron particles are deposited at the leading edge of the heat source; as the heat moves over these particles, they are sintered into a layered, glassy deposit. By contrast, the PCVD process employs direct Radio Frequency (RF) excitation of a microwave-generated plasma. Because the microwave field can be moved very quickly along the tube (since it heats the plasma directly, not the silica tube itself), it is possible to traverse the tube thousands of times and deposit very thin layers at each pass, which makes for very precise control of the preform index profile. A separate step is then required for sintering of the glass. In both cases, the preforms require a final heating to around 2150°C in a furnace to collapse the preform into a state from which glass is ready to be drawn. All inside vapor deposition (IVD) processes require that a tube be used as a preform; minor flaws in the tube can induce corresponding dips and peaks in the fiber index profile.

In the so-called outside process, a rotating, thin cylindrical mandrel is used as the substrate for subsequent CVD; the mandrel is then removed before the preform is sintered. An external torch fed by carrier gasses is used to supply the chemical components for the reaction, as well as to provide the necessary heat for the reaction to occur. Two outside processes that have been widely used are the outside vapor deposition (OVD) and the vapor axial deposition (VAD) methods. OVD was among the first procedures developed; it is basically a flame hydrolysis process in which the torch consists of discrete holes in a pattern of concentric rings, each of which provides a different constitutent element for the chemical reactions. The VAD process is similar in concept, using a set of concentric annular apertures in the torch; in this case, the preform is pulled slowly across the stationary torch. By mixing $GeCL_4$ as a dopant into the $SiCl_4$—O_2 feed, the proportion of germania (GeO_2) deposited with the silica varies with the

temperature of the flame; if a wide flame is used, the temperature gradient produces a graded portion of germania deposit.

Silica glass has absorption bands in both the ultraviolet (UV) and mid-infrared (IR) wavelength ranges, which provide a fundamental limit to the attenuation that can be achieved. This occurs despite the fact that the Raleigh scattering contribution decreases inversely as the fourth power of the wavelength, and the UV Urbach absorption edge decreases even faster with increasing wavelength. The infrared absorption increases at long wavelengths, becoming dominant at wavelengths greater than about 1.6 microns, which results in the minimal loss for wavelengths around 1.55 microns. Note that for many years, optical fiber attenuation was limited by a strong hydroxide (OH) absorption band near wavelengths of 1.4 microns; this has been steadily reduced over time with improved fabrication methods, until the loss minimum near 1.55 microns was brought close to the Raleigh limit. The best dopants for altering the refractive index of silica glass are those that provide a weak index change without inducing a large shift in the UV absorption edge. These include GeO_2 and P_2O_5 (which increase the refractive index in the fiber core) and B_2O_3 and SiF_4 (which decrease the refractive index in the cladding). The chemical reactions for the general process often use high vapor pressure liquids such as $GeCl_4$, $POCl_3$, or SIF_4 along with oxygen as a carrier gas; these reactions are well documented [2–4]. It is more difficult to introduce more exotic dopants, such as the rare-earth elements used in optical fiber amplifiers, and there is not a single widely used technique at this time. There are also other preform fabrication and fiber drawing techniques that are not generally used for telecommunication grade optical fiber, but can be important for other material systems and applications. These include bulk casting of the preform and the non-CVD method of "rod and tube" casting (in which the core and cladding are cast separately and combined in a final melting step). There are also preform-free drawing techniques such as the "double crucible" method, in which the core and cladding are formed separately in a pair of platinum crucibles and combined in the drawing process itself. This method was important in the past, but is not widely used today because it does not provide the same precision control as the drawing tower process.

A great deal of research has been done into fiber materials with better transparency in the infrared. Although none of these materials has yet proven to be a serious competitor with doped silica, they may prove useful for other applications that do not have the strict requirements of telecommunication systems, such as CO_2 laser transmission, medical applications, or remote sensing and imaging [5]. For example, bulk infrared optics and fibers can be made from sulfide, selenide, and telluride glasses; collectively known as chalcogenide fibers, they exhibit transmission loss on the order of 1 dB/meter in the wavelength range 5–7 microns [6]. Another example, the heavy-metal fluroide fibers [7], holds some interest for communication systems because their theoretical limit for Rayleigh scattering is

much lower than for silica (this is due to a high-energy ultraviolet absorption edge and better infrared transparency). However, excess absorption has proven difficult to reduce, and current state-of-the-art fluroide fibers continue to exhibit losses near 1 dB/km, well above the Rayleigh limit. Fabrication of short lengths of fluroide fiber has been somewhat more successful in transmitting longer wavelengths, with reported losses as low as 0.025 dB/km at wavelengths of 2.55 microns [8]. The residual loss mechanisms in longer fibers remain the subject of ongoing research and are thought to be due to extrinsic impurities or defects such as platinum particles from the fabrication crucibles, bubbles in the core preform and at the core-cladding boundry, and fluoride microcrystals.

Many types of optical fiber are available for different environments, including undersea cables, both outdoors (with integrated strength members to facilitate hanging cables from telephone poles) and indoors. The properties of optical fiber cables are governed by various industry standards, as noted in Chapter 1, including G.652 (with a maximum bandwidth of 50 GHz) and G.655 (with a maximum bandwidth of several THz). As discussed in Chapter 7, fiber jackets are typically rated as either riser, plenum, low smoke, low/zero halogen, or a dual-rated combination of the above; the IEC 1034 specification provides the dual-rated jacket specifications, while the zero halogen jackets are typically free of chlorine, fluroine, and bromides. New cable jacket types are also emerging from the national fire and electrical codes, such as the "limited combustability" designation for some types of plenum cables. Some cable types are rated for use under a computer room raised floor, and others for installation via air blowers in plenum ducts. As discussed in Chapter 2, many types of multifiber connectors, structured cabling systems, and cable pullers, conduits, and patch panels are available. Special designs of conventional optical connectors are available, such as the so-called elite MT connectors, a customized version of the 12-fiber MT ferrule that is sorted to guarantee less than 0.35 dB maximum loss per fiber. New types of multifiber connectors are also being developed using two-dimensional ferrules to accommodate 24 or more fibers in a single plugging operation. In addition, many companies manufacture custom optical fibers to a user-specified refractive index profile, doping, core or cladding geometry, numerical aperture, cutoff wavelength, or other characteristics. Short sections of optical fiber may be packaged as loopbacks or wrap plugs for transceiver testing, while long haul spools may require special shielding for mechanical or structural reasons. Specialty fibers with coatings to increase their mechanical performance under bending are used today in small form factor VF-45 style connectors; these are described in more detail in Chapter 3. Furthermore, there are many types of fiber-optic connectors available, including so-called no polish field-installable connectors, which are intended for quick and low-cost installation by untrained personnel, physical and nonphysical contact connectors, flat ferrules and angle-polished ferrules for low back reflectance, connectors with built-in variable attenuators or mode scramblers, metal-

lized fibers which are soldered into place, and many, many other variants. While it is not practical to give a comprehensive list of every specialty fiber type in this chapter, we will provide an overview of some major fiber types that are commonly encountered, as well as a synopsis of emerging fiber types that may become important in the future.

Many simple devices such as fiber-optic splitters, couplers, and combiners can be manufactured; the most common techniques include fiber tapering and fusion splicing [11–14], etching [15], and polishing [16–18]. One of the most common devices is a tapered fiber-optic power splitter, often implemented in single-mode fiber [19]. In this process, two glass fibers with their protective jackets removed are brought close together and parallel to each other, then fused and stretched using a heat source. Light that is initially launched into only one fiber will be initially transferred to the cladding interface as it enters the tapered region, then to the core-cladding mode of the adjacent fiber, and finally back to the core mode of the second fiber (this is known as a cladding mode coupling device [20, 21]). In some cases, such as optical power splitters, it is more desirable to remove the dependence of coupled power on wavelength; acromatic couplers can be fabricated by using two fibers with different propagation constants. These are known as dissimilar fibers; in most cases, fibers are made dissimilar by changing their cladding diameters or cladding indecies. In this case, the above equation for coupled power must be modified and the power vs. distance is not simply sinusoidal, but becomes much more complex [22]. Other approaches are also possible, such as tapering the device such that the modes expand well beyond the cladding boundaries [23], or encapsulating the fibers in a third material with a different refractive index [24–26].

4.2. NEXT-GENERATION OM3 MULTIMODE FIBER

Conventional datacom links use single-mode fiber for long-distance, high-speed links and multimode fiber for shorter links. Early datacom applications, including ESCON, Token Ring, FDDI, Ethernet, and ATM, operated at relatively slow data rates (4–155 Mbit/s), using low-cost infrared light-emitting diode transmitters (LEDs). The earliest fibers, called Optical Multimode 1 (OM1), featured a larger core than is used today and a bigger numerical aperture. As the technology matured, smaller core multimode fiber was typically rated for a minimum bandwidth-distance product around 160 MHz-km for 62.5/125 micron fiber at 850-nm wavelength; 500 MHz-km for 50/125 micron fiber at this wavelength; and 500 MHz-km for both fiber types at 1300-nm wavelength. This fiber was compatible with various industry standards, including CCITT recommendation G.652 (see Chapter 2), and was defined by the ISO standards as “optical multimode 2” (OM2) fiber; it is also commonly known as “FDDI grade” fiber. The fiber bandwidth was measured using an overfilled launch (OFL) test procedure, which

replicated the large spot size and uniform power profile of an LED. Since an LED consistently fills the entire fiber core, the fiber bandwidth is determined by the aggregate performance of all the excited modes (see Chapter 1 or 12 for calculations showing the number of modes in a fiber). However, LED sources typically have a maximum modulation rate of a few hundred Mbit/s; with the growing demand for higher data rates, laser sources operating over single-mode fiber were required (see Chapter 5).

Single-mode links using Fabry-Perot or distributed feedback lasers operating at long wavelengths (1300 nm) tend to be higher cost due to their tighter alignment tolerances and higher performance characteristics. There is a lower cost alternative; the recent deployment of short-wave (780–850 nm) vertical cavity surface emitting lasers (VCSELs) has made it possible to use multimode fiber at higher data rates over longer distances. Compared with LEDs, VCSELs offer higher optical power, narrower spectral width, smaller spot size, less uniform power profiles, and higher modulation data rates. This means that a VCSEL will not excite all of the modes in a multimode fiber; the fiber bandwidth is determined by a restricted set of modes, typically concentrated near the center of the core. Older multimode fibers experienced significant, often unpredictable variations in bandwidth when used with VCSEL sources due to defects or refractive index variations in the fiber core and variations in the number and power of excited modes due to fluctuations in the VCSEL output or between different VCSEL transmitters.

In response to these problems, the datacom industry developed a new type of laser-optimized or laser-enhanced multimode fiber specifically designed to achieve improved, more reliable performance with VCSELs. Precise control of the refractive index profile minimizes modal dispersion and differential mode delay (DMD) with laser sources, while remaining backward compatible with LED sources (the dimensions, attenuation, and termination methods for laser-optimized and conventional fiber are the same). The first laser-optimized fibers, introduced in the mid-1990s, were available in both 50-micron and 62.5-micron varieties and designed for 1-Gbit/s operation up to a few hundred meters. These fibers were not always capable of scaling to higher data rates; with the increased attention on 10-Gbit/s links, newer types of laser-optimized fiber were developed, which further enhanced the fiber core profile. While some types of 62.5-micron laser-optimized fiber were capable of reaching about 35 meters at 10 Gbit/s, it became apparent that the smaller core diameter and reduced number of modes in 50 micron fiber made it the preferred choice for these data rates. Today, laser-optimized fiber is commonly available only in 50-micron versions, with an effective bandwidth-distance product around 2000 MHz-km for 850-nm laser sources. The bandwidth must be measured using a restricted mode launch (RML) test, instead of the conventional OFL method. (For more details, refer to the Telecommunication Industry Association (TIA) task group on modal dependencies of

bandwidth, TIA FO-2.2.) This fiber was defined in the TIA-568 standard as "laser-optimized multimode fiber," and in the ISO 11801 (2nd edition) by its more common name, "optical multimode 3" (OM3) fiber.

An early example of laser-optimized fiber is the Systimax LazrSPEED[1] fiber introduced by Lucent, which uses a green jacket to distinguish it from existing multimode (orange), single-mode (yellow), and dispersion-managed (purple) fiber cables. Attenuation is about 3.5 dB/km at 850 nm and 1.5 dB/km at 1300 nm; bandwidth is 2200 MHz-km at 850 nm (500 MHz-km overfilled) and 500 MHz-km at 1300 nm (no change when overfilled). Another example is the Corning InfiniCore[2] fiber, which typically uses an aqua-colored cable; the CL 1000 line consists of 62.5-micron fiber made with an outside vapor deposition process that achieves 500-m distances at 850 nm and 1 km at 1300 nm. Similarly, the CL 2000 line of 50-micron fiber supports 600-m distance at 850 nm and 2 km at 1300 nm.

Most recent installations of Ethernet, Fibre Channel, InfiniBand, and other systems use the preferred OM3 fiber, and many legacy systems including ESCON are compatible with this fiber. In order to avoid the cost associated with installing new fiber, most standards attempt to accommodate various types of multimode fiber. While the idea of backward compatibility works reasonably well up to 1 Gbit/s (distances of a few hundred meters can be achieved), it begins to break down at higher data rates when the achievable distance is reduced even further. Designing a future-proof cable infrastructure under these conditions becomes increasingly difficult; at some point, new fiber needs to replace the legacy multimode fiber. Although single-mode fiber should be a good long-term investment, the short-term cost premium for single-mode fiber installation and ports on many switches, servers, and storage devices remains a concern. Since the cost of short-wave transceivers is presently lower than long-wave transceivers, there is still some question as to the preferred fiber to install and the best mixture of 62.5-micron and 50-micron multimode fiber. In general, 50-micron fiber has been widely deployed in Europe and Japan, while North America has primarily used 62.5-micron multimode fiber until recently. The IEEE has recommended using 62.5-micron multimode fiber in building backbones for distances up to 100 m, and 50-micron fiber for distances between 100 and 300 m. Mixing OM2 and OM3 fibers in the same link results in an aggregate bandwidth proportional to the weighted average of the two cable types (see Case Study on Multimode Fiber). Care must be taken not to mix 50- and 62.5-micron fibers in the same cable plant, as the resulting mismatch in core size and numerical aperture creates high losses. This can make it difficult to administer a mixed cable plant, as there is no industry standard connector keying to prevent misplugging different types of multimode fiber into the wrong location.

---

[1]Systimax and LazrSPEED are trademarks of Lucent Corporation.

[2]InfiniCore CL 1000 and CL 2000 are trademarks of Corning Corporation.

4.3. OPTICAL MODE CONDITIONERS

Because of the bandwidth limitations of multimode optical fiber, future multi-gigabit fiber-optic interconnects will be based on single-mode fiber cables. For this reason, most new fiber installations include at least some single-mode fiber in the cable infrastructure. However, many applications continue to use multimode fiber extensively; a recent survey of building premise cable installers reported that most LAN infrastructures currently installed are composed of about 90% multimode fiber [32]. As the fiber cable plant is upgraded to support higher data rates on single-mode fiber, we must also provide a migration path that continues to reuse the installed multimode cable plant for as long as possible. The need to migrate from multimode to single-mode fiber affects many important datacom applications [33]:

- I/O applications currently using multimode fiber for ESCON will need to migrate the cable plant to single-mode fiber in order to take full advantage of the higher bandwidth of FICON links. Future FICON enhancements that extend this protocol to multi-gigabit data rates will also require single-mode fiber.
- Networking applications such as ATM have traditionally used different adapter cards to support multimode and single-mode fiber. The gigabit Ethernet standard (IEEE 802.3z) is the first industry standard to propose the use of both fiber types with the same adapter card.
- Parallel Sysplex links were originally offered as either 50-Mbyte/s data rates over multimode fiber or 100-Mbyte/s data rates over single-mode fiber. With the announcement of more recent servers, support for multimode fiber has been withdrawn as a standard feature and is now available only on special request. There is a need to support 100-Mbyte/s adapter cards over installed multimode fiber to facilitiate migration of those customers who have been using the 50-Mbyte/s option.

In order to address these concerns, special fiber-optic adapter cables have been developed known as mode conditioning patch cables (MCP). These cables contain both single-mode and multimode fibers, and should be inserted on both ends of a link to interface between a single-mode adapter card and a multimode cable plant. This allows the maximum achievable distance for multimode fiber (550 m) and enables some applications to continue using the installed multimode cable plant. The MCPs for parallel sysplex links, Gigabit Ethernet, Fibre Channel, and many other applications are available today.

Next, let us describe the technical issues associated with this approach. The bandwidth of an optical fiber is typically measured using an overfilled launch condition, which results in equal optical power being launched into all fiber modes [32]. This is also known as a mode scrambled launch and is approximately

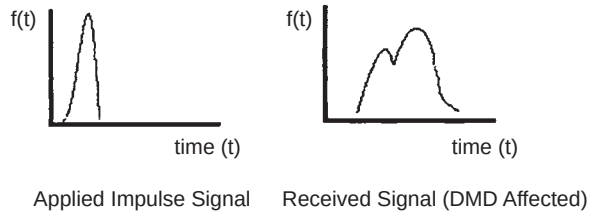


Figure 4.1 Effect of DMD on signals passing through multimode fiber.

equivalent to the conditions achieved when using a Lambertian source such as an LED. By contrast, laser sources, being more highly collimated, tend to produce an underfilled launch condition; this can result in either larger or smaller effective bandwidth relative to an overfilled launch, and is sensitive to small changes in the fiber's refractive index profile. As discovered in recent gigabit link tests [34], bandwidth measured using overfilled launch conditions is not always a good indication of link performance for laser applications over multimode fiber. As illustrated in Fig. 4.1, when a fast rise-time laser pulse is applied to multimode fiber, significant pulse broadening occurs due to the difference in propagation times of different modes within the fiber. This pulse broadening is known as differential mode delay (DMD); it is observed as an additional contribution to timing jitter (measured in ps/m) and can be large enough to render a gigabit link inoperable. DMD values are unique to the modal weighting of a source, the modal delay and mode group separation properties of the fiber, mode-specific attenuation in the fiber, and the launch conditions of the test. DMD is made worse by the excitation of relatively few modes with similar power levels in widely spaced mode groups and a high percentage of modal power concentrated in lower order modes. The impact of DMD increases with link length. Unfortunately, there is not a simple relationship between the industry specified overfill launch measured bandwidths of the fiber and the effective bandwidth due to DMD.

The radial overfill launch method was developed as a way to establish consistent and repeatable modal bandwidth measurement of a given fiber coupled with a given source [34]. A radial overfill launch is obtained when a laser spot is projected onto the core of the multimode fiber, symmetric about the core center with the optic axis of the source and fiber aligned; the laser spot must be larger than the fiber core, and the laser divergence angle must be less than the fiber's numerical aperture. When these conditions are satisfied, the worst case modal bandwidth of the link is taken to be the worse of the overfill and radial overfill launch condition measurements (although for most applications, the radial overfill launch will be the worst case). There is a good correlation between the radial overfill launch bandwidth and the DMD limited bandwidth of a fiber. Thus, high-speed laser

links implemented over multimode fiber will likely experience bandwidth values closer to the radial overfill launch method rather than the more commonly specified overfill launch method.

To allow for laser transmitters to operate at gigabit rates over multimode fiber without being unduly limited by DMD, a special type of fiber-optic jumper cable was developed to "condition" the laser launch and obtain an effective bandwidth closer to that measured by the overfill launch method. The intent is to excite a large number of modes in the fiber, weighted in the mode groups that are highly excited by overfill launch conditions, and to avoid exciting widely separated mode groups with similar power levels. This is accomplished by launching the laser light into a conventional single-mode fiber, then coupling into a multimode fiber that is off-center relative to the single-mode core, as shown in Fig. 4.2. The offset launch can be introduced in two ways. One version requires manufacturing a splice between the single-mode and multimode fiber with a controlled amount of lateral offset between the fiber cores. A tolerance analysis of this approach revealed that some installations could experience unacceptable variability in the splice elements, resulting in poor alignment and ineffective mode conditioning. For this and other reasons, the preferred embodiment uses standard ceramic ferrule technology with an offset in the ferrule alignment. Different offsets are required for 50.0- and 62.5-micron multimode fiber cores. Evaluations conducted by the Gigabit Ethernet Task Force, Modal Bandwidth Investigation Group, have verified that single-mode to 62.5-micron multimode MCPs with lateral offsets in

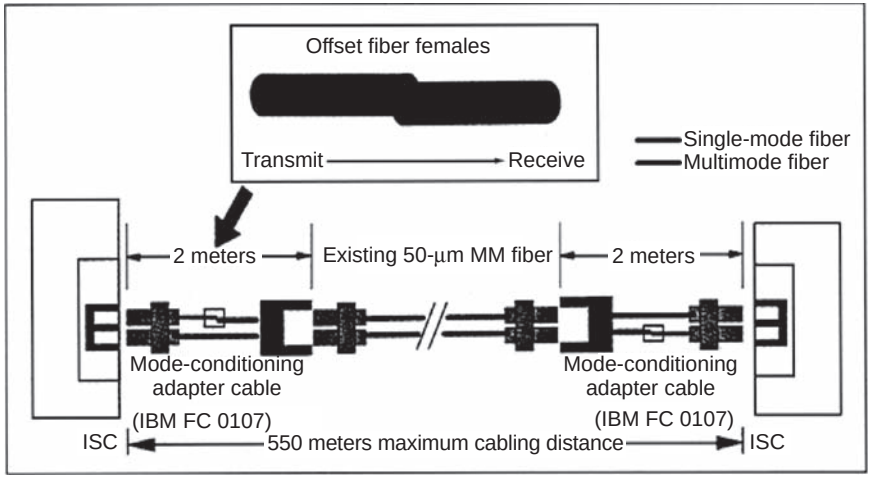


Figure 4.2 Off-center ferrule design for mode conditioning patch cables.

the 17–23 micron range can achieve an effective modal bandwidth equivalent to the overfill launch method across 99% of the installed multimode fiber infrastructure. Similar work has shown that single-mode to 50-micron multimode offset launch cables with lateral offsets in the 10–16 micron range will achieve similar results. The MCP is similar to a standard 2-m jumper cable, except that it contains both single-mode and multimode fibers and includes a small package for the offset ferrules near one end (32, 33). Alternate designs for MCPs have also been proposed, in which the offset launch condition is replaced by special optics that convert the laser spot into a donut-shaped launch.

4.4. ATTENUATED FIBER CABLES FOR WDM AND CABLE TV

Some applications require optical attenuators to control or limit the optical power on a fiber link. One example is wavelength division multiplexing (WDM) equipment, as described in Chapter 15, which uses fixed attenuators to allow a common transceiver to interoperate with many different physical layers. Other examples include fiber links in the cable television industry. Fixed attenuators can be expensive and must be incorporated into the design of the cable system; this means that duplex cables cannot be used if attenuation is required in only one path of a duplex fiber link. Separating duplex connectors defeats the keying, which prevents the connector from being improperly inserted into a receptacle or transceiver. The attenuators also provide an extra connection point in the link, which must be cleaned and may be susceptible to mechanical vibration that will tend to dislodge connectors in data communication products. Instead of using fixed, pluggable attenuators, it is possible to manufacture inline optical attenuators as part of the fiber-optic cable assembly. Several approaches can be used. For multimode attenuators, a short piece of single-mode fiber can be spliced into the cable; by controlling the alignment between the single-mode fiber stub and the multimode cable on either side, as well as the length of the stub, various levels of attenuation can be achieved. The stub may be actively aligned during cable maufacturing, then protected by an external sheath or package to protect it from mechanical shock, vibration, and cable flexing. Similar effects can be achieved by deliberate misalignment of a multimode or single-mode fiber fusion splice. In most cases, arbitrary attenuation values from 0.5 dB to over 20 dB can be realized with a tolerance of less than 0.5 dB. The resulting attenuated cables are quite robust, and in many cases they serve a dual purpose since the application would have required jumper cables to adapt to different styles of optical connectors or connect with subtended equipment.

For other applications in which optical power must be controlled, specialty fibers are available with high attenuation that is flat over a certain spectral region.

With so much design effort directed toward reducing the fiber attenuation to improve link budgets and distances, it can be easy to forget that optical fiber can just as easily be designed for high attenuation. Using the same precision-controlled manufacturing techniques that consistently yield low-loss fiber, it is possible to dope the fiber in such a way that very consistent, high attenuation is provided over a wide range of operating wavelengths. This can be an advantage in designing attenuators for WDM systems. As with the offset spliced cable attenuators, these fibers tend to be more reliable and robust than conventional airgap attenuators. They have the additional advantage that a controlled amount of attenuation can be selected by simply cutting and splicing a desired length of the fiber. Short sections of high-attenuation fiber are also being integrated in some types of fiber-optic components and are finding applications in optical test and measurement systems. Typical fibers are available with attenuations ranging from 0.5 dB/meter to 30 dB/meter in 0.5-dB increments, or from 0.25 dB/cm to 25 dB/cm in 0.25-dB increments.

4.5. POLARIZATION CONTROLLING FIBERS

In the design of fiber-optic systems it is important to know how many modes can propagate in the fiber, the phase constants of the different modes, and their spatial profiles. To do this we need to solve the wave equation for a particular fiber geometry as described in Chapter 1. The solution depends on the specific refractive index profile of the fiber. For the case of step-index fiber profiles, a complete set of analytical solutions exists [35]; these can be grouped into three different types of modes depending on the direction of the electric field vector relative to the direction of propagation. They are called transverse electric (TE), transverse magnetic (TM), and hybrid modes. The hybrid modes can be further separated into two classes depending on whether the electric field, E, or magnetic field, H, is largest in the transverse direction; these are called EH and HE modes, respectively. In practice, the refractive index difference between the core and cladding of an optical fiber is so small (about 0.002 to 0.009) that most of these modes are degenerate and it is sufficient to use a single notation for all modes, called the linearly polarized or LP notation. An LP mode is denoted by two subscripts, which refer to the radial and azimuthal zeros of the particular mode; for example, the fundamental mode is the LP_{01} mode. This is the only mode that will propagate in a single-mode fiber.

The cylindrical symmetry of an optical fiber leads to a natural decoupling of the radial and tangential components of the electric field vector. Hence, standard single-mode fiber does not maintain the polarization state of the light when it is launched. However, these two polarizations are nearly degenerate, and a fiber with circular symmetry is most often described in terms of orthogonal linear polarizations. This near-degeneracy of the two polarization modes is easily broken

by any imperfections in the cylindrical symmetric geometry of the fiber, including mechanical stresses on the fibers. These effects can either be introduced intentionally during the fiber manufacturing process or inadvertently after the optical fiber has been installed. The effect is known as birefringence; it results in two orthogonally polarized modes with slightly different propagation constants (note that the two modes need not be linearly polarized, and in general they will be elliptical polarizations). Because each mode experiences a slightly different refractive index, the modes will drift in phase relative to each other; at any point in time, the light in the fiber exists in a state of polarization that is a superposition of the two orthogonally polarized modes. Birefringence of a fiber may be specified as the difference in refractive index between the two modes of propagation. The net polarization evolves as the light propagates through various states of ellipticity and orientation. After some distance, the two modes will differ in phase by a multiple of 2π, resulting in a state of polarization identical to that at the fiber input. This characteristic length is known as the beat length, and is a measure of the intrinsic material birefringence in the fiber; the time delay between the two modes is called polarization dispersion, and it can impact the performance of communication links in a manner similar to intermodal dispersion [9]. For example, if the time delay is less than the coherence time of the light source, the light in the fiber remains coherent and fully polarized. For sources of wide spectral width, however, this condition is reversed and light emerges from the fiber in a partially polarized and unpolarized state (the orthogonal polarizations have little or no statistical correlation). Links producing an unpolarized output can experience a 3-dB power penalty when passing through a polarizing optical element at the output of the fiber.

A stable polarization state can be ensured by deliberatly introducing birefringence into an optical fiber; this is known as polarization preserving fiber or polarization maintaining fiber (PMF). Fibers with an asymmetric core profile will be strongly birefringent, having a different refractive index and group velocity for the two orthogonal polarizations (this is sometimes known as loss discrimination between modes). Such fibers are useful in some types of systems that require control of the transmitted light polarization. There are many possible core configurations, as shown in Fig. 4.3; for example, elliptical cores provide a simple form of PMF by using very high levels of dopant in the core. These so-called high-birefringence fibers also experience high attenuation because of the elevated dopant levels in the core. A double-core geometry (not shown in the figure) will also introduce a large birefringence. Another approach is to create mechanical stress within the fiber, such as in the bow tie configuration, by inserting stress inducing members near the fiber core. Note that in all of these examples, polarization is preserved only if the initial signal is polarized along one of the preferred directions in the PMF; otherwise, the polarization of the signal will continue to drift as light propagates along the fiber.

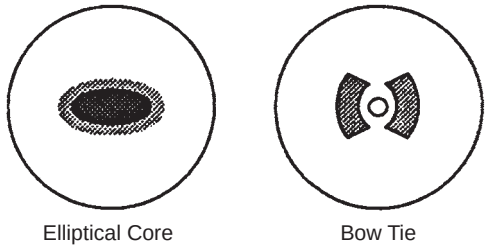


Figure 4.3 Cross sections of polarization maintaining fibers.

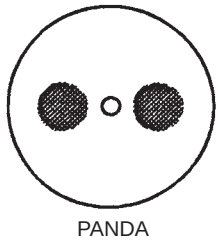


Figure 4.4 Cross-section of PANDA fiber, showing mechanical members that apply strain to the fiber.

Another example is the Polarization Maintaining and Absorption Reducing (PANDA) fiber; the areas highlighted in Fig. 4.4 show parts of the fiber core doped to create an area with a different coefficient of expansion than that of the cladding. In manufacturing, as this fiber cools, stresses are set up due to this dfference, which in turn modifies the refractive index without requiring high levels of dopants in the core. These stress-applying members are present along the entire length of the fiber; such fiber is commercially available.

There are other ways to make PMF, although they are not widely used in commercial products [36]. For example, so-called low-birefringence fibers can be made by very carefully controlling the fiber profile, since there is no reason for power to couple between orthogonally polarized modes if there are no irregularities in a perfectly circular fiber. Another way is to twist the fiber during manufacturing, or deliberately make the core off-center, so that the two polarization modes become circular in opposite directions and power coupling cannot take place. Yet another variation is called spun fiber; a PANDA fiber preform is spun while the fiber is being drawn, producing a full revolution about every 5 mm. Spun fiber has no polarization dependence at all, but is very difficult to make successfully at lengths much beyond about 200 m and is very expensive. Consequently, it is not commonly used for communication systems.

Polarization effects can manifest themselves in a number of ways. For example, standard erbium doped fiber amplifiers (EDFAs) exhibit two forms of birefringence, which are usually considered trivial but may build up in systems with many optical amplifiers. First, polarization-dependent loss (PDL) refers to the fact that most EDFAs exhibit higher gain for one polarization state than for the orthogonal state. Since the arriving signal is composed of a superposition of states, the gain changes slowly (over a timescale of minutes); however, the amplified spontaneous emission noise is unpolarized and experiences a fixed gain. Hence, there is variation in the signal-to-noise ratio over time. Note that this effect accumulates as the root mean square of all the amps in a chain, rather than as a straight summation. Second, polarization-dependent gain (PDG) is a saturation effect in the EDFA itself; the amplifier exhibits a higher gain in the polarization state orthogonal to that of the signal. One way to combat these effects is by polarization scrambling of the input signals. Another is to design a polarization insensitive EDFA; this can be done by using a circulator to direct light into a Faraday polarization rotating mirror such that the light makes two trips through the EDFA. The PDL and PDG effects introduced on one polarization state in the first pass through the EDFA are also induced in the second polarization state during the second pass through the EDFA, such that the emerging light has uniform gain across both polarization states.

Recently, new types of polarization-maintaining fibers have been announced for high-capacity fiber systems. The Lucent TruePhase family of polarization-maintaining fibers is offered in application-specific wavelengths. A new nonzero-dispersion fiber from Pirelli, the Advanced FreeLight fiber, features a reduction of 50% in its polarization mode dispersion ratio over previously available fibers. As this is a rapidly evolving field, we can expect many new fiber types to be introduced in the coming years with improved properties.

4.6. DISPERSION CONTROLLING FIBERS

As discussed in Chapter 1, multimode optical fibers are subject to modal dispersion, whereas both multimode and single-mode fibers experience a combination of material (or chromatic) dispersion and waveguide dispersion. It was also noted that chromatic and waveguide dispersion have opposite signs so they may cancel each other out; this is why conventional silica fiber has a dispersion minima around 1300-nm wavelength. We may group together the collective effects of all these factors under the term *group velocity dispersion* (GVD). Standard single-mode fiber can exhibit either normal or anomalous dispersion. Under normal dispersion, long wavelengths have a higher group velocity than short wavelengths; if a wide spectrum of light is launched into this fiber, the red wavelengths will emerge first, followed by the blue wavelengths (this is also known as a positive frequency chirp or up-chirp). For anomalous dispersion, the situation is

reversed; short wavelengths travel faster than long wavelengths, and blue light will emerge from the fiber before red light (this is called a negative frequency chirp or down-chirp).[3] Since modulation of a signal necessarily increases its bandwidth, and all practical light sources have some finite spectral width, dispersion effects occur in all communication systems.

As shown earlier, a typical silica optical fiber has an attenuation minimum around wavelengths of 1.55 microns and a zero-dispersion point at 1.3 microns. This represented a fundamental tradeoff in fiber link design, depending on whether the optical links were intended to be loss limited or dispersion limited. Shorter distance data communication systems such as ESCON typically chose to operate at 1.3 microns, while links designed for long-haul communications were designed around 1.55 microns. Both the loss minimum and material dispersion are inherent physical properties of the silica fiber materials. However, waveguide dispersion can be affected by the refractive index profile design [36]. The profile shown in Fig. 4.5 has been successfully used to shift the zero-dispersion point 1.55 microns; this is known as dispersion-shifted fiber (DSF). Conventional fiber that has not been treated in this manner is called nonzero-dispersion shifted fiber (NZDSF). Currently available DSF has a number of practical problems. It is more prone to some forms of signal nonlinearity, especially becuase its slightly smaller mode field diameter concentrated electromagnetic field more strongly in the core. WDM systems can also experience strong interchannel interference, or so-called near end crosstalk (NEXT). For example, nonlinear effects such as four-wave mixing (FWM), also known as four-photon mixing, can be a serious design issue for WDM systems (see Chapter 15). FWM is strongly influenced by the wavelength channel spacing and by the fiber dispersion. In order for FWM to occur, each channel must stay in phase with its adjacent channel for a considerable distance. Thus, if fiber dispersion is high (as with standard NDSF in the 1550-nm band,

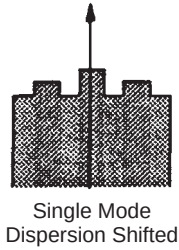


Figure 4.5 Index profile of dispersion-shifted fiber.

[3]Note that while this terminology is consistent with most other reference books, in some engineering texts the meaning is reversed; the definition given here for normal dispersion is called anomalous, and the definition given here for anomalous is called normal.

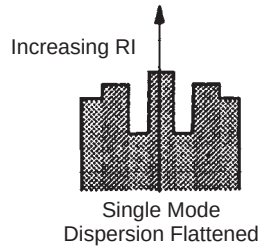


Figure 4.6 Index profile of dispersion-flattened fiber.

typically around 17 ps-nm-km) the effects of FWM are minimal for channel spacings greater than about 25 Ghz (if channel spacing is reduced below about 15 GHz, the effect of FWM can be severe even on standard fiber). If DSF is used (dispersion less than 1 ps-nm-km), then FWM effects are maximized; the effect causes degradation at channel spacings of less than 80 GHz, and at a channel spacing of 25 Ghz around 80% of the optical energy in the original two signals will be transferred into either sum or difference frequencies. So, one might ask why all fiber is not dispersion shifted to take advantage of the minimal dispersion properties. Part of the answer is that DSF significantly increases problems like FWM. Other nonlinear effects, such as stimulated Raman scattering (see Chapter 7), are also worse when operating over DSF.

In order to address these problems, new types of fiber are available which essentially guarantee a certain level of dispersion (around 4 ps-nm-km), although these are currently very expensive. Other variants known as dispersion-optimized fiber are also available, with a refractive index profile as shown in Fig. 4.6. This guarantees around 4 ps-nm-km dispersion in the 1530–1570 nm wavelength range; it is available under various brand names, including Tru-Wave[4] fiber from AT&T (nonzero-dispersion-shifted) and so-called SMF-LS fibers from Corning.

In addition, more sophisticated dispersion compensation can be achieved by adding several core and cladding layers to the fiber design. The index profile shown in Fig. 4.7 is quite complex, but it is possible to realize dispersion less than 3 ps-nm-km over the entire wavelength range 1300 nm–1700 nm using this approach. This is known as dispersion-flattened fiber; it was intended to allow users to easily migrate from 1300-nm systems to 1550-nm systems without changing the installed fiber. This fiber also has potential applications to broadband WDM applications, for which fiber dispersion must be kept as uniform as possible over a range of operating wavelengths. The principal drawback is its high loss, around 2 dB/km, which prevents general use in the wide area network.

[4]Tru-Wave is a trademark of AT&T.

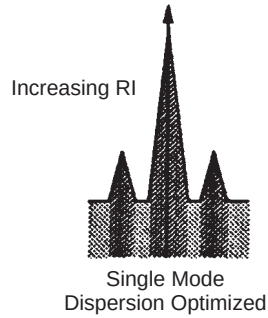


Figure 4.7 Index profile of dispersion-optimized fiber.

It is also possible to construct a fiber index profile for which the total dispersion is over 100 ps-nm-km in the opposite direction to the material dispersion; this can be used to reverse the effects of conventional fiber dispersion. So-called dispersion-compensating fiber (DCF) is commercially available with an attenuation of around 0.5 dB/km. DCF has a much narrower core than standard single-mode fiber, which accentuates nonlinear effects; it is also typically birefringent and suffers from polarization mode dispersion.

Different fiber designs have recently been introduced that are intended for use in WDM environments. For example, Corning has recently introduced its large effective area fiber, known as LEAF, for use in the 1550-nm window (both C-band and L-band wavelengths) at data rates up to 10 Gbit/s and beyond. The LEAF fiber is a single-mode, nonzero-dispersion shifted fiber with a larger effective area in which optical power can be transmitted. Typically, this fiber can accommodate 2 dB more than conventional fibers without introducing nonlinear effects that can arise because of high-power levels in the core, espeically at the high-power levels associated with multiwavelength WDM systems. LEAF fiber has enhanced bandwidth and effectively quadruples the information carrying capacity of the fiber. This has made LEAF fiber particularly well suited to long-distance carriers and datacom service providers. A variation on this technology is the so-called MetroCor[5] fiber from Corning, a single-mode NZDSF compatible with industry standard G.655 that is designed to handle both C-band and L-band transmission in metropolitan area networks (1280 to 1625 nm). MetroCor is an example of so-called negative dispersion fiber, which is expected to improve performance of dense wavelength division multiplexing (DWDM) systems; it has

[5]MetroCor is a trademark of Corning Corporation.

recently been demonstrated as part of a DWDM metropolitan area network testbed. Similarly, the AllWave[6] fiber from Lucent provides a 50% larger spectrum than conventional fiber, lowering the attenuation between the 1300-nm and 1550-nm windows while maintaining low dispersion at 1300 nm. To protect against bending loss on single-mode WDM systems, special fiber is available such as the Blue Tiger[7] cables from Lucent. Identified by a blue cable jacket, this fiber is specially designed to support very tight bend radius applications (only 0.3-dB loss on a 10-mm bend, as compared with 1 to 1.5 dB for a conventional fiber under these conditions). Since a sharp fiber bend can induce enough loss to force WDM equipment to protection switch, this type of fiber may be important as WDM finds increasing applications in the metropolitan area network.

Dispersion-flattened fibers have recently been investigated in soliton propagation systems for long-distance communication without optical repeaters. We mention this for completeness, since datacom systems typically do not use soliton transmission; this is presently reserved for long-haul telecommunication links. The nonzero-dispersion slope in the fiber causes different wavelengths to experience different average dispersions; this can be a significant limitation in classical soliton long-distance WDM transmission. By combining different types of commercially available fiber with different signs of dispersion and dispersion slopes, it is possible to design paths with essentially zero-dispersion slope. For example, this was demonstrated in a recent experiment in which almost flat average dispersion (D = 0.3 ps/nm-km) was achieved by combining standard, dispersion compensated, and Tru-Wave fibers; this enabled soliton transmission of 27 WDM channels, each carrying 10 Gbit/s, over more than 9000 km. Soliton experiments may lead to other types of specialty fibers; for example, adiabatic soliton compression can be obtained by using a fiber whose dispersion slowly decreases with distance (so-called dispersion tapered fiber).

4.7. PHOTOSENSITIVE FIBERS

Many types of glass are sensitive to ultraviolet light, which can induce a permanent change in their refractive index. These are known as photosensitive or photorefractive materials; in this case, the refractive index profile of the fiber can be changed by light that either propagates along the fiber length or illuminates an unjacketed fiber from the side. Many different types of materials can be used for this effect [37]; standard glass fiber can be made photosensitive by doping it with hydrogen, for example, while other fiber types do not require the hydrogenation process. If the fiber is illuminated through a transmisssion mask, or by an

[6]AllWave is a trademark of Lucent.
[7]Blue Tiger is a trademark of Lucent.

interference pattern created by two light beams, then the photorefractive effect can be used to write a diffraction grating into the fiber index variations. Recently, a class of devices known as inline fiber Bragg gratings (FBGs) have been developed that hold great promise for many applications such as optical filters, wavelength add/drop multiplexers, and dispersion compensators. Writing a fiber Bragg grating requires a larger full-width at half maximum response due to saturation of the index change. Due to coupling of radiation modes, undesirable side lobes of the grating can also be created at shorter wavelengths. However, this can be controlled with more recently developed types of fiber and grating writing techniques. For example, a grating stronger than 30 dB can be written in a few minutes using specialty fibers, with side lobes kept below 0.1 dB.

Photosensitive fibers can also be used to write chirped fiber Bragg gratings for dispersion compensation. Recently, it has been demonstrated that the delay curve of grating-based dispersion compensators can be designed for a nearly perfectly linear response. This has caused increased interest in the use of FBGs for dispersion compensation in WDM systems. The FBG does have a certain amount of intrinsic polarization mode dispersion, but this can be overcome by coupling them with a suitable PMD compensator. This type of dispersion compensation is one of the many alternatives to dispersion shifter, flattened, and compensated fibers discussed earlier. With the increased interest in high bit rates (10 to 40 Gbit/s and beyond) over long-haul distances (hundreds of km), the market for dispersion compensation is growing rapidly, and there will likely be many types of specialized optical fiber designed for different applications.

4.8. ACTIVE OPTICAL CABLES

Due to the inherent bandwidth-distance limitations of high-speed copper cable assemblies, it is not uncommon for designers to embed amplifiers, signal regenerators, or similar devices within a cable assembly. These components are powered by voltage carried along the cable, either through dedicated wires or by repurposing extra ground wires. These so-called active cables or “bump in the wire” approaches can double the achievable unrepeated distance, although they do add cost compared with passive cables. Recently, there has been increased interest in active optical cable assemblies for datacom applications. Rather than use a separate optical transceiver and pluggable fiber cable, this approach embeds the active transceiver components into the connector and permanently attaches the optical fiber at either end. The result is a cable assembly that appears to have standard electrical connectors on both ends; the optical design is hidden from the end user. As with active electrical cables, extra ground pins in the connector must be repurposed to provide power for the laser, receiver, and related circuitry. Unlike an active copper cable, the active components are located near the cable ends, so voltage does not have to be transmitted a significant distance.

This approach has many potential advantages. Compared with parallel copper wire (such as CX4), parallel active optical cables are significantly lighter weight and more flexible. If the network equipment is properly designed, it becomes possible to have a common electrical port that can accept either passive copper cable or active copper cable, allowing the user to select the appropriate technology and cost with easy reconfiguration. The optical connector can be an expensive, precision element (particularly for parallel optical links) due to tight manufacturing tolerances. This is eliminated by permanently attaching the fibers to the optics. Contamination and cleaning of the optical interface as well as connector loss repeatability are also eliminated. Furthermore, an active cable allows the manufacturer to control both ends of the link, and removes interoperability standards, eye safety, and other constraints on the optical interface. For example, it becomes possible to pair a high-power transmitter with a sensitive receiver, knowing they will never be separated in the field. This should reduce costs associated with interoperability testing and improve manufacturing yields by allowing reuse of components outside normal specification limits. These factors are expected to allow active optical cables to be cost competitive with copper at much shorter distances than previously possible, and could lead to displacement of copper by optical links in some applications. Of course, other technical challenges remain for this emerging technology. The strain relief between the cable ends and fiber becomes a critical part of this design. Connectors that contain active components must be ruggedized to withstand the harsh environment within a data center. The active components are located just outside the server package; this might lower the cooling burden on the server or might make the optics harder to cool and pose reliability issues. Active optical cables must be stocked in different lengths, just like copper cables. Reliability of the active optical cables must also be addressed, as with any new technology. As of this writing, several companies are offering active optical cables; while they are not yet standardized, they are allowed by the most recent issue of the InfiniBand standard. Consequently, many active optical cables feature an InfiniBand copper interface; 4X is the most popular, although 12X versions are sampling or under development. Initial per-fiber data rates are in the 2.5- to 5-Gbit/s range, with higher rates anticipated in the future. Many clustered servers and blade servers are considering active optical cables as a cost-effective alternative to conventional transceivers and cable assemblies.

REFERENCES

1. Nolan, D. 2000. Tapered fiber couplers, mux and demux. Chapter 8 in *Handbook of optics*, vols. III and IV. Washington, D.C.: Optical Society of America.
2. Hill, K. 2000. Fiber Bragg gratings. Chapter 9 in *Handbook of optics*, vol. IV. Washington, D.C.: Optical Society of America.
3. Adams, M. J. 1981. *An introduction to optical waveguides*. Chichester: John Wiley and Sons.

4. Carlisle, A. W. 1985. Small size high performance lightguide connectors for LANs. *Proc. Opt. Fiber Comm. Paper TUQ* 18:74–75.
5. Miller, S. E., and A. G. Chynoweth. 1979. *Optical fiber telecommunications*. New York: Academic Press.
6. Nishii, J., et al. 1992. Recent advances and trends in chalcogenide glass fiber technology: A review, *Jour. Non-crystalline Solids* 140:199–208.
7. Sanghera, J. S., B. B. Harbison, and I. D. Agrawal. 1992. Challenges in obtaining low loss fluoride glass fibers. *Jour. Non-crystalline Solids* 140:146–149.
8. Takahashi, S. 1992. Prospects for ultra-low loss using fluoride glass optical fiber. *Jour. Non-Crystalline Solids* 140:172–178.
9. DeCusatis, C. 2001. Design and engineering of fiber optic systems. From *The Optical Engineer's Desk Reference*, ed. E. Wolfe. Washington, D.C.: Optical Society of America.
10. Hect, E., and A. Zajac. 1979. *Optics*, New York: Addison-Wesley.
11. Nolan, D. 2000. Tapered fiber couplers, mux and demux. In *Handbook of optics*, vol. IV, Chapter 8. McGraw-Hill, New York, NY.
12. Ozeki, T., and B. S. Kawaski. 1976. New star coupler compatible with singe multimode fiber links. *Elec. Lett.* 12:151–152.
13. Kawaski, B. S., and K. O. Hill. 1977. Low loss access coupler for multimode optical fiber distribution networks. *App. Opt.* 16:1794–1795.
14. Rawson, G. E., and M. D. Bailey. 1975. Bitaper star couplers with up to 100 fiber channels. *Elec. Lett.* 15:432–433.
15. Sheem, S. K., and T. G. Giallorenzi. 1979. Singlemode fiber optical power divided; encapsulated teching technique. *Opt. Lett.* 4:31.
16. Tsujimoto, Y., et al. 1978. Fabrication of low loss 3 dB couplers with multimode optical fibers. *Elec. Lett.* 14:157–158.
17. Bergh, R. A., G. Kotler, and H. J. Shaw. 1980. Singlemode fiber optic directional coupler. *Ele. Lett.* 16:260–261.
18. Parriaux, O., S. Gidon, and A. Kuznetsov. 1981. Distributed coupler on polished singlemode fiber. *App. Opt.* 20:2420–2423.
19. Kawaski, B. S., K. O. Hill, and R. G. Lamont. 1981. Biconical-tapered singlemode fiber coupler. *Opt. Lett.* 6:327.
20. Lamont, R. G., D. C. Johnson, and K. O. Hill. 1984. *App. Opt.* 24:327–332.
21. Snyder, A., and J. D. Love. 1983. *Optical waveguide theory*, New York: Chapman and Hall.
22. Brown, T. 2000. Optical fibers and fiber optic communications. In *Handbook of optics*, vols. III and IV, Chapter 1. McGraw-Hill, New York, NY.
23. Weidman, D. L. 1993, December. Achromat overclad coupler. U.S. Patent 5,268,979.
24. Truesdale, C. M., and D. A. Nolan. 1986. Core-clad mode coupling in a new three-index structure. European Conf. on Optical Comm., Barcelona, Spain.
25. Keck, D. B., A. J. Morrow, D. A. Nolan, and D. A. Thompson. 1989. *Jour. Lightwave Tech.* 7:1623–1633.
26. Miller, W. J., D. A. Nolan, and G. E. Williams. 1991, April 1. Method of making a 1 X N coupler. U.S. Patent no. 5,017,206.
27. Fiber transport services physical and configuration planning (IBM document number GA22-7234). 1998. Mechanicsburg, Pa.: IBM Corp.
28. Planning for fiber optic channel links. 1993. (IBM document number GA23-0367). Mechanicsburg, Pa.: IBM Corp.
29. Maintenance information for fiber optic channel links. 1993. (IBM document number SY27-2597). Mechanicsburg, Pa.: IBM Corp.
30. DeCusatis, C. 1998, December. Fiber optic data communication: Overview and future directions. *Optical Engineering*, special issue on Optical Data Communication.

31. See Lucent product information at www.lucent.com.
32. Giles, T., J. Fox, and A. MacGregor. 1988. Bandwidth reduction in gigabit Ethernet transmission over multimode fiber and recovery through laser transmitter mode conditioning. *Optical Engineering* 37:3156–3161.
33. DeCusatis, C., D. Stigliani, W. Mostowy, M. Lewis, D. Petersen, and N. Dhondy. 1999, September/November. Fiber optic interconnects for the IBM Generation 5 parallel enterprise server. *IBM Journal of Research and Development* 43, no. 5/6:807–828.
34. Giles, C. R. 1997. Lightwave applications of fibre Bragg gratings. *IEEE Journ. Lightwave Tech.* 15, no. 8:1391–1404.
35. Okoshi, T. 1982. *Optical fibers.* New York: Academic Press.
36. Dutton, H. 1999. *Optical communications,* New York: Academic Press.

This page intentionally left blank

Case Study: Multimode Fiber Reuse for High-Speed Storage Area Networks

Many studies have attempted to characterize the existing installed optical fiber cable infrastructure to determine whether it is suitable for 10-Gbit/s applications. For example, a study presented to the IEEE 802.3 group and cited by the ANSI T11 committee provides detailed analysis of installed fiber broken down for different parts of the global market [1]. The estimated distribution of optical fiber link distances within building backbones is shown in Fig. 1; the installed length distributions are in general agreement with previous studies, although there is some variance in the details (for example, while there is general agreement that over 85% of installed links are less than 300 m, estimates of the breakdown in link distances between 100 m and 300 m can vary by a factor of two compared with the results shown in this figure). For our purposes, it is sufficient to note that there are a significant number of installed optical links at 100 m or less and a significant fraction of low bandwidth fiber in use for these applications.

Many enterprise Storage Area Networks (SANs) use a combination of short jumper cables, patch panels, and fiber trunks covering distances on the order of 100 m under a computer center raised floor, within the walls or plenum air ducts, or in an overhead conduit. An enterprise data center may have over 25,000 patch panel connections with associated optical cables. To reduce cost and minimize disruption to the network, these cables are typically reused for as long as possible, spanning many generations of servers and storage devices; it is easier to relocate patch panels and use longer jumper cables than to make changes to the installed trunk cables. In fact, it is assumed that installed optical fiber in building backbones is never replaced; rather, new fiber is simply installed in the same conduit as the old fiber, which was intended to support FDDI or ESCON. It is a testament to the high reliability of fiber-optic cables that most of this infrastructure remains

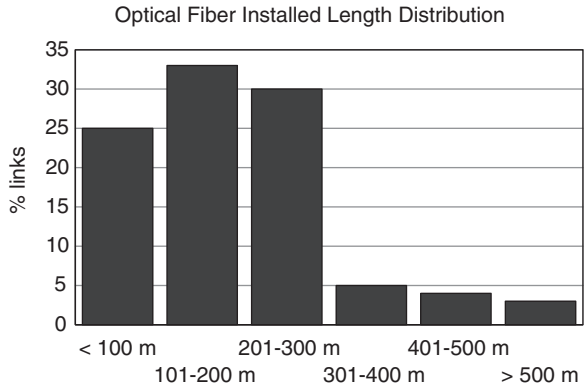


Figure 1 Estimated distribution of installed optical fibers (from Ref. [1]).

in use today. Thus, it has become a significant, practical problem to determine how long the installed multimode fiber infrastructure can continue to be reused.

Historically, several attempts have been made to accommodate intermixing of multimode fiber types, and to support legacy multimode fiber using single-mode transceivers. For example, ESCON allowed the use of 62.5-micron jumper cables on either end of a 50-micron trunk, provided that there was no more than one transition from 62.5-micron cores to 50-micron cores on a unidirectional link segment (see Chapter 21). Another notable effort to reuse multimode fiber involved the use of mode conditioners, or mode conditioning patch cables (MCPs), to interconnect single-mode long-wavelength laser transceivers with legacy multimode fiber (see Chapter 4). This approach was adopted by early gigabit link designs, including Ethernet, Fibre Channel, and FICON, to facilitate migration from legacy multimode fiber. Using mode conditioners on either end of the link, it was not necessary to replace both the fiber and transceivers (server adapter cards) at the same time. Instead, the adapter cards could be changed immediately, while the upgrade from multimode to single-mode fiber could be deferred to a later time, provided that the links were fairly short (less than about 300–550 m, depending on the protocol). While this was a reliable technical solution, MCP cables were fairly expensive to produce, and the approach was not extendable to higher data rates; thus, there are no provisions for MCPs in 2 Gbit/s Fibre Channel or higher data rate standards. As data rates increase, the achievable maximum link distance is, of course, reduced; for example, the maximum distances supported at different data rates for different fiber types is shown in Fig. 2 (although these distances are for Fibre Channel SAN links, the specifications for Ethernet are almost identical).

The requirement for higher data rates, combined with a desire to reuse installed fiber, has led to increased interest in using OM3 fibers in combination with legacy

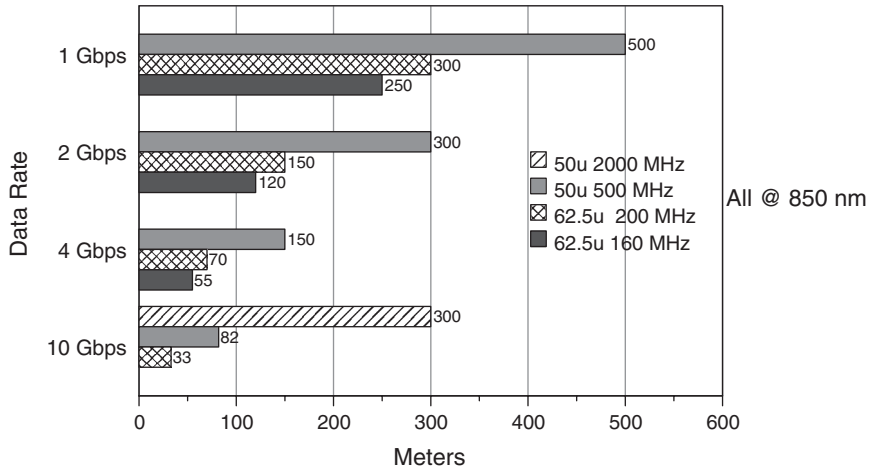


Figure 2 Reducation in multimode fiber bandwidth-distance product.

multimode cables. Most manufacturers do not recommend mixing different types of multimode fiber in the same link. This may seem nonintuitive at first; if 50-micron fiber has improved bandwidth compared with 62.5-micron fiber, then one might expect that replacing at least some of the installed fiber would provide at least a partial benefit. This is not the case; in fact, mixing different fiber types can lead to instabilities in the link because of technical issues that arise when coupling one fiber type to another. This problem has been widely described in the technical literature [2–4] and is not supported by any of the recent industry standards. When coupling from a smaller diameter fiber into a larger one, there is very little effect; the larger fiber acts as a bucket to capture the entire optical signal. However, all data communication links are bidirectional, so each transition from smaller to larger diameter fiber must also include a corresponding transition from larger to smaller diameter fiber. This causes two significant effects. First, as one might expect, some of the optical signal is lost during the transition to smaller core fiber. Second, and more importantly, not all of the propagating modes in the larger core fiber will be supported in the smaller core fiber. This leads to modal noise effects, which depend on the type of laser source, the laser encircled flux, modal power distribution, and spot size, and the fiber's bandwidth and numerical aperture. Simulation and measurements on thousands of links [3] make it possible to characterize the impulse response of the fiber; these calculations suggest that mode mixing can lead to a high statistical probability that the link will fail at some point during its lifetime.

If we limit ourselves to the effects of combining fiber of different bandwidths into a single link, however, we can estimate the effective distances that can be

achieved to a good approximation by assuming Gaussian-shaped transfer functions for the fiber sections. This simplified approach should be used only as a planning guideline, since it does not account for statistical variations in the properties of specific fibers. For example, consider a previously installed 50-micron multimode cable with a low bandwidth, B1 (500 MHz-km). The original cable was 75 m long; it is now required to reach a new piece of equipment that is 100 m away, so a new section of cable is attached. In an effort to achieve the best possible performance, the new cable uses more recently available fiber with a higher bandwidth, B2 (2000 MHz-km). We are asked to determine whether this combination can support transmission of a 10-GHz signal.

Since the fiber core diameter is the same for both the trunk and jumper cables, and the refractive index profile is almost the same, there will be negligible excess loss due to coupling or numerical aperture mismatch between the high and low bandwidth cables. We can approximate the overall link bandwidth, B, by adding the reciprocal bandwidths of each fiber segment in the link, in much the same way as electronic resistors are added in parallel ($1/B = 1/B1 + 1/B2$). The bandwidth of the first link segment, B1, is (500 MHz-km/0.75 km) = 6666.66 MHz, while the bandwidth of the second link segment, B2, is (2000 MHz-km/0.25 km) = 80,000 MHz. Adding these as reciprocals yields B = 6153.84 MHz for the total 100-m link, or 615.38 MHz-km effective bandwidth-distance product. We note that transmission of a 10-GHz signal over this link is only possible up to a maximum distance of (615.38 MHz-km/0.1 km) = 61.5 m. Therefore it is not possible to operate this link in the desired configuration. There is too much low-bandwidth fiber; we need to reduce either the link distance or the data rate to achieve a working configuration. It may be necessary to install new higher bandwidth fiber for the entire link, rather than attempting to reuse the installed fiber, in this example. Similar calculations may be performed for different combinations of fiber lengths and bandwidths, higher data rates, or more than two concatenated fiber sections.

REFERENCES

1. Flatman, A. 2004, March. In-premises optical fibre installed base analysis to 2007, comissioned by Agilent & Cisco Systems, presented to IEEE 802.3 10 Gigabit over FDDI-grade Fibre study group meeting, Orlando, Fl. Also available from www.ieee.org/standards
2. DeCusatis, C. 2005, March. Fiber optic cable infrastructure and dispersion compensation for storage area networks, *IEEE Communications Magazine*, 43, no. 3:86–92.
3. Pepeljugoski, P., J. Abbott, and J. Tatus, "Effect of launch conditions on power penalties in gigabit links using 62.5 micron core fiber operating at short wavelength, *IEEE Transactions*, to be published.
4. Pepeljugoski, P., J. Schaub, J. Tierno, J. Kash, S. Gowda, B. Wilson, H. Wu, and A. Hajimiri. 2003. Improved performance of 10 Gb/s multimode fiber optic links using equalization. Paper THG4, *Proc. OFC* 2003, 2:472–474.

5

Optical Sources: Light-Emitting Diodes and Laser Technology

Wenbin Jiang
Michael S. Lebby
Phoenix Applied Research Center, Motorola, Incorporated, Tempe, Arizona 85284

5.1. INTRODUCTION

In this chapter, we will present some fundamentals of optical transmitter design for fiber-optic data communications. Specifically, we will consider the operating principles behind the three most common types of optical sources used for data communications:

- Light-emitting diodes (LEDs)—edge emitting (E-LEDs) and surface emitting
- Edge-emitting semiconductor laser diodes
- Vertical cavity surface-emitting lasers (VCSELs)

As discussed in Chapter 2, the sources used for fiber-optic communications most often operate at wavelengths near 1300 or 1550 nm, and VCSELS are currently limited to operation near 780–980 nm. A detailed discussion of laser and LED sources could easily fill a book by itself; for our purposes, we will present some of the fundamentals of this technology and refer the interested reader to some of the many good treatments in the literature [1–3]. Fabrication and manufacturing of these devices, with particular emphasis on VCSELs, will be discussed further in Chapter 9.

5.2. TECHNOLOGY FUNDAMENTALS

An optical source for fiber-optic communications must have properties that are slightly different from those required for other applications. Specifically, it must exhibit a high radiance over a narrow band of wavelengths that coincide with the transmission window of the fiber, typically 0.8–1.55 μm. The emissive area should

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

be no greater than the core diameter of the fiber, and the distribution of incident radiation should match the numerical aperture of the fiber (the acceptance cone). In addition, the source should be easily modulated at high frequencies (up to several gigahertz), have a modest cost compared with other components of the datacom system, and be at least as reliable as other large system computer components (several hundred thousand power-on hours with minimal variation in optical output power). Semiconductor sources satisfy these requirements by emission of light at a pn junction through the process of injection luminescence. In this chapter, we will begin with a brief review of semiconductor properties and then describe the fundamentals of LEDs, edge-emitting laser diodes, and VCSELs.

For the range of wavelengths commonly used in data communications, the efficiency of both sources and detectors is determined by the bandgap energy. It is often desirable for the bandgap of the detectors to be slightly less than that of the source materials. To begin, then, we will consider some fundamentals of semiconductor physics (more information on this topic is provided in Chapter 6). Intrinsic semiconductor materials are characterized by an electrical conductivity much lower than that of pure metals, which increases rapidly with temperature. As described in Chapter 6, the Fermi level lies between the conduction and valence bands of the material. These materials may be doped with various impurities as either *n* type or *p* type, so that the Fermi level moves closer either to the conduction or valence bands. Common room-temperature semiconductors include germanium and silicon, from group IV(b) of the periodic table, as well as several binary compounds from groups III(b) and V(b), such as gallium arsenide (GaAs) and indium phosphide(InP).

Silicon or germanium can be made *n* type by doping with donor impurities, such as P, As, or other group V elements, whereas they can be made *p* type by doping with acceptor impurities such as B, Ga, or other group III elements. These dopant concentrations are fairly small, approximately le-14 to le-21 per cubic centimeter in Si—or approximately from 1 dopant atom per billion silicon atoms to approximately 1 per thousand silicon.

Typically, the Fermi level is uniform throughout the doped material; when an abrupt junction between *p*-type and *n*-type materials is formed, the conduction, valence, and Fermi bands bend to accommodate the difference. Electrons tend to accumulate on the *n* side of the junction, and holes develop on the *p* side. A thin depletion region is formed at the junction itself—the depletion region is effectively depleted of both hole and electron carriers.

An external voltage applied across the junction will have the effect of raising or lowering the potential barrier between the two materials, depending on its polarity. When the *p* side is connected to a positive voltage, we say that the junction is forward biased; the opposite condition is known as reverse bias. In a forward-biased junction, excess carriers are injected into either side of the depletion

region, where they are the minority carriers. These excess minority carriers tend to diffuse away from either side of the depletion layer. Thickness of the depletion layer decreases under forward bias conditions. There is thus a balance of carrier flow; injection of carriers over the depletion region (which gives rise to an excess carrier concentration) is balanced by both diffusion away from the junction and carrier recombination.

Recombination is the process in which the electrons and holes combine. Several different kinds of recombination are possible, although we are only concerned with the radiative recombination processes that release photons: these are direct band-to-band transitions across the bandgap. Direct bandgap semiconductor materials (such as GaAs and other III–V compounds) have an energy–*k* relationship such that the conduction band minimum and the valence band maximum occur at the same value of *k*, as shown in Fig. 5.1. With direct bandgap materials, the holes and electrons can recombine directly; by contrast, in an indirect bandgap material such as silicon, the conduction band minimum does not correspond to the valence band maximum, and it is more complicated to recombine while satisfying the requirements of both conservation of energy and conservation of momentum. The nonradiative transitions, in which the energy is lost to thermal

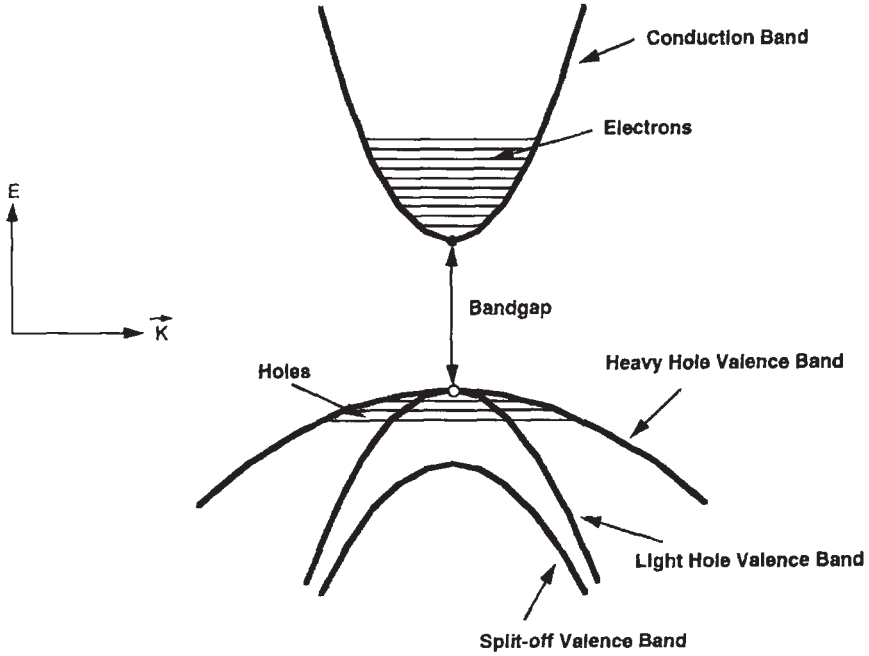


Figure 5.1 Energy band structure (energy vs *k*) for direct bandgap semiconductor.

energy within the crystal, involve intermediate energy levels that trap carriers deep within the crystal lattice.

In direct bandgap materials, the radiative recombination process is proportional to the excess minority carrier concentration and gives rise to the creation of radiation from the junction. The process is known as injection luminescence and is the fundamental mechanism for the creation of optical radiation from semiconductor optical sources. We will not discuss this effect in great detail, other than to note that the operation of such sources is also influenced by the minority carrier lifetime, diffusion length near the junction, injection efficiency, and width of the depletion layer. Equivalent circuits for the pn junction (discussed in Chapter 3) allow the modeling of junction capacitance and series resistance; these factors affect how rapidly the optical source can be modulated or, conversely, how fast a semiconductor photodiode can respond to incident radiation.

The recombination of electrons and holes in a forward-biased pn junction can produce optical radiation at wavelengths of 0.8–1.7 μm, suitable for data communications. If recombination takes place in several stages, it is also possible for more than one photon of longer wavelength to be emitted. When only a single photon of energy (E) is produced, the wavelength may be determined from Planck's Law:

$$\lambda = hc/E = 1.24/E(\mu m/eV) \tag{5.1}$$

where h is Planck's constant and c is the speed of light. The probability that an electron of energy E_2 will recombine with a hole of energy E_1 is proportional to the concentration of electrons at E_2 and the concentration of holes at E_1. The spectrum of the radiated energy may be determined by integrating the product of these carrier concentrations over all values of E_1 or E_2, subject to the constraint that the difference between E_1 and E_2 is the desired photon energy. Using a simple model for the behavior of carriers in a semiconductor, it is possible to derive the emission spectra due to injection luminescence; an LED will give rise to an approximately Gaussian spectra with a half-width of between 30 and 150 nm or more at room temperature.

It is desirable to design devices with high internal quantum efficiency, defined as the ratio of the rate of photon generation to the rate at which carriers are injected across the junction. A related parameter is the external quantum efficiency, defined as the ratio of the number of emitted photons to the number of carriers crossing the junction. External quantum efficiency is smaller than internal quantum efficiency for several reasons. For instance, some light will be reabsorbed before it can reach the emitting surface. Only light emitted toward the semiconductor–air interface is useful for coupling into the fiber, and a small percentage of this light is reflected back from the surface. Furthermore, only light reaching the surface at less than the critical angle will be coupled into an adjacent optical

fiber. For a semiconductor material with an internal quantum efficiency of η_i, an injection current of I will give rise to an optical power, P, of

$$P = \eta_i \cdot (I/q) \cdot E \tag{5.2}$$

where q is the charge on an electron. It is also possible to derive expressions for the behavior of these sources at high modulation frequencies; it can be shown that they exhibit a high-frequency cutoff at modulation frequencies greater than

$$f = 1/\tau_p \tag{5.3}$$

where τ_p is the mean lifetime of the excess minority carriers; it gives rise to a characteristic diffusion length, L,

$$L = \sqrt{\mu \cdot k \cdot T \cdot \tau_p / q} = \sqrt{(D \cdot \tau_p)} \tag{5.4}$$

where μ is the electron mobility, k is the Boltzmann constant, and the quantity D is known as the electron diffusion coefficient. From these expressions, we can determine that modulation frequencies above f = 20–25 MHz are difficult to obtain. However, bandwidths as high as several hundred MHz can be obtained from the double-heterostructure diodes that we will discuss in this chapter.

The structure of a typical surface-emitting LED is shown in Fig. 5.2. Note that for some structures, part of the substrate may be etched away to minimize the distance between the active area and the emitting surface. A thin layer of insulating oxide may be used to separate the positive contact from the *p* layer, except in the region of the active layer where the current flow is concentrated. The LED junction temperature is a critical parameter; the wavelength distribution of the LED changes with increasing temperature, internal quantum efficiency falls off, and the lifetime of the device is significantly decreased. In general, for GaAs and GaAlAs devices, the peak junction temperature should be kept below 50 or 100°C by the use of heat sinks or other devices. If the active layer is kept close to the heat sink, the thermal impedance is small, and higher current densities may be used without an excessive rise in temperature.

In general, LEDs are diffuse (Lambertian) sources; for a typical GaAs LED forward biased at 100 mA and consuming 150 mW of power, the optical output is only on the order of 50 μW. To compensate for this rather low power, various

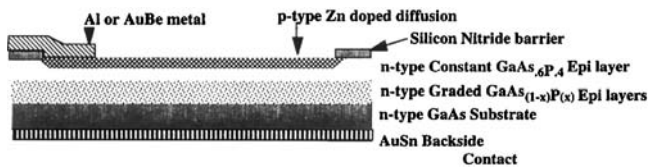


Figure 5.2 Cross-sectional view of an LED.

optical coupling and packaging schemes are employed; these will be discussed in other chapters. One alternative is the edge-emitting LED, which offers small emissive area and higher radiance. The edge-emitting LED takes advantage of a slightly different approach—the double-heterojunction diode. Heterojunctions are produced at the boundary between two different types of materials whose physical lattice sizes match but that have different bandgap energies and other properties. Without going into details about the physics involved, we note that such structures offer higher injection efficiency, improved carrier confinement (and hence higher probability for radiative transitions to occur), and higher modulation bandwidths. The double heterostructure is simply a layered approach that further helps to confine excess minority carriers that are injected over the forward-biased pn junction. For such devices, the modulation bandwidth increases as the square root of the current density and inversely as the square root of the active layer thickness (note that the modulation current must be a small fraction of the DC bias current, or nonlinear effects will occur that limit device performance). With current densities of 50 A/mm^2, devices such as this routinely offer 1- or 2-mW output power from a 50-μm aperture. These can also be made to emit at the longer wavelengths—in the 1.3- to 1.55-μm range.

The light output of an LED is based on spontaneous emission and tends to be incoherent, with typical spectral widths on the order of 100–150 nm. The two types of LEDs, surface-emitting LEDs and ELEDs, are characterized by high divergence of the output optical beams (120° or more), slow rise times (>1 ns) that generally limit the data communication applications of LEDs to less than 200 MHz, low temperature dependence, insensitivity to optical reflections, and output powers on the order of 0.1–3 mW.

By contrast, laser sources tend to exhibit much higher output powers (3–100 mW or more), narrower spectral widths (<10 nm), smaller beam divergence (5–l0°), faster modulation rates (hundreds of MHz to several GHz), and higher sensitivity to both temperature fluctuations and optical reflections back into the laser cavity. This is because they are based on stimulated emission of coherent radiation—indeed, the word "laser" is an acronym for light amplification by stimulated emission of radiation. The radiative recombination process responsible for injection luminescence in LEDs is the result of spontaneous bandgap emissions. Transitions may also be stimulated by the presence of radiation of the proper wavelength; all photons produced by stimulated emission have the same frequency and are in phase with the stimulating emission, making this a source of coherent radiation. This is the type commonly produced by semiconductor edge-emitting lasers and vertical cavity lasers. Like all lasers, the semiconductor laser diode requires three essential operating elements:

1. A means for optical feedback, usually provided by a cleaved facet or multilayer Fresnel reflector

2. A gain medium, consisting of a properly doped semiconductor material
3. A pumping source, provided by current applied to the diode

For a semiconductor material with gain g and loss α, placed in a cavity of length L between two partially reflective mirrors of reflectivity R_1 and R_2, the minimum requirement for lasing action to occur is that the light intensity after one complete hip through the cavity must at least be equal to its starting intensity. This is called the threshold condition, given by the following relation:

$$R_1R_2\exp(2(g - \alpha)L) = 1 \tag{5.5}$$

Thus, the laser diode begins to operate when the internal gain exceeds a threshold value. Two key parameters associated with the laser are its efficiency of converting electrical current into laser light, known as the slope efficiency *dP/dI*, and the amount of current required before the laser begins stimulated emission, known as the threshold current I_{th}. This information is summarized in the power vs. current (*P* vs. *I*) characteristic curve of a laser diode, as shown in Fig. 5.3. Applying current, *I*, to a laser device initially gives rise to spontaneous emission of light, as in an LED; as the current is increased, the cavity and mirror losses are overcome, and the laser passes the threshold and begins to emit coherent light. The optical power emitted above threshold, *P*, is given by

$$P = (I - I_{th})dP/dI \tag{5.6}$$

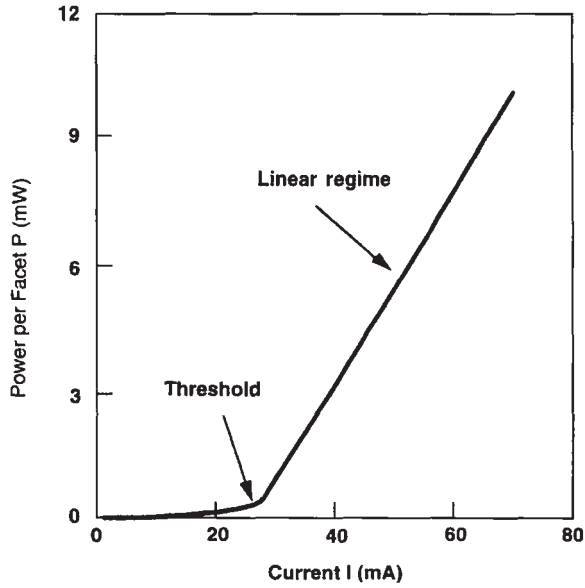


Figure 5.3 Power vs. current (*P* vs. *I*) characteristic curve of a laser diode.

Threshold current is found by extrapolating the linear region of the *P/I* curve. Low-threshold currents and high efficiencies are desirable in laser diodes. The slope efficiency, *dP/dI*, is related to the external differential quantum efficiency, η, by

$$\eta = (\lambda q)/(hc)(dP/dI) \tag{5.7}$$

where λ is the wavelength, *q* is the charge of an electron, *h* is the Plank constant, and *c* is the speed of light. The threshold current is a strong function of temperature, *T*, according to the relation

$$I_{th}(T) = I_{th}(T_1)\exp[(T - T_1)/T_0] \tag{5.8}$$

where T_0 is the laser's characteristic temperature, and T_1 is room temperature (300 K). A large T_0 is desirable because it translates into a reduced sensitivity of threshold current to operating temperature. Usually, it is only possible to measure the case temperature of a packaged laser diode. The junction temperature is always higher than the case temperature in typical applications. The junction temperature can be estimated by taking into account the device thermal resistance, the input electrical power, and the output optical power.

Increasing the diode current above threshold allows the gain coefficient to stabilize at a value near the threshold point, just large enough to overcome losses in the media:

$$g = \alpha - \frac{1}{2L}\ln(R_1R_2) \tag{5.9}$$

Because of the high gain coefficients that can be obtained, the cavity of a double heterostructure laser can be made much smaller than other types of lasers; 1 mm or less is typical. For lasers biased above threshold, the slope of the power vs. current characteristic curve in the spontaneous emission region corresponds approximately to the external quantum efficiency. The slope in the lasing region is related to the differential quantum efficiency, ηd, by

$$\eta_d = \frac{q}{E}dP/dI \tag{5.10}$$

In practice, this ideal characteristic curve may be less well behaved; considerable research has been devoted to eliminating undesirable effects, such as "kinked" curves that correspond to unwanted fluctuations in the laser power. This has led to the development of many features, such as stripe geometry and buried heterostructure lasers, which are beyond the scope of this discussion. Compared with LED emissions, laser light has a narrower bandwidth (<10 nm, making it less susceptible to certain kinds of noise in the fiber link), it is more directional (so that the external quantum efficiency may be improved), and its modulation bandwidth is greater (up to several Ghz). Laser sources at short wavelengths (780–

850 nm) are typically used with multimode fiber, whereas longer wavelengths (1.3 or 1.55 μm) are used with single-mode fiber.

5.3. DEVICE STRUCTURE—LED

This and the following sections will introduce device structures for LEDs, edge-emitting lasers, and VCSELSs, with particular emphasis on the laser technologies. In Chapter 6, the manufacturing issues for each will be introduced, again with particular emphasis on the laser technologies.

LEDs use either direct bandgap or indirect bandgap materials for spontaneous light emission. When holes, in the valence band of Fig. 5.1, combine with electrons in the conduction band, photons of light with energy consistent with the bandgap are emitted. The material selection is dependent on the desired wavelength of light to be emitted. For example, whereas GaA emits light at approximately 880 nm, GaP emits light at approximately 550 nm and InGaP emits light at approximately 670 nm. The light emission begins as soon as the LED is forward biased: Unlike the laser diodes discussed in the following section, there is no threshold current.

A surface-emitting LED consists of a pn junction, with a controlled concentration profile between the *p*-typed and *n*-typed material. Figure 5.2 shows the cross section of an LED. The substrate in this example is GaAs; other III–V compounds, such as GaP, may be appropriate depending on the desired wavelength of light to be emitted. The substrate layer and the graded layer above are doped *n* type, and they are also doped to provide the bandgap for the desired wavelength of emitted light. A constant layer caps the graded layer and is capped in turn by a dielectric material such as silicon nitride. The *p* region of the pn junction is formed by the diffusion of Zn doping into an opening in the dielectric. The pn junction is contacted on the top by a metal such as gold, AuBe, or aluminum contacting the *p*-type, diffused Zn-doped region. The *n*-type region is contacted on the back side, with a metal such as AuSn or AuGe in contact with the *n*-doped substrate.

Device characteristic curves for LEDs include the forward current vs. applied voltage curve and the light emission or luminous intensity vs. forward current curve. Other useful characteristics in optoelectronic applications include the spectral distribution and the luminous intensity versus temperature and angle, the latter due to the need to couple into a medium such as optical fiber. An example of each is shown in Fig. 5.4. The spectral distribution of an LED is wider than that of a laser, incurring more chromatic dispersion. The lower frequency with which an LED can be switched, the wider angle of emission, and the larger spectral distribution are all disadvantages of LEDs compared to lasers. The primary advantage of LEDs over lasers for data communication is the low cost due to the significantly simpler processing.

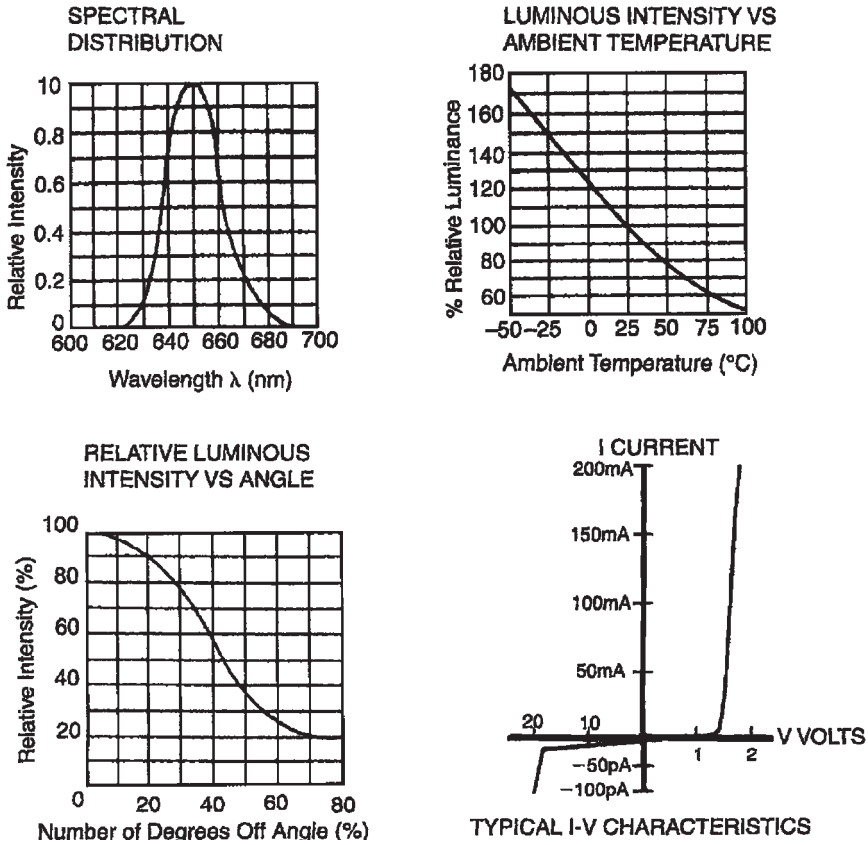


Figure 5.4 Characteristics of an LED: spectral distribution, luminous intensity versus temperature and angle, and I versus V characteristics.

5.4. DEVICE STRUCTURE—LASERS

Edge-emitting semiconductor lasers have been around for more than 30 years. Semiconductor lasers were first reported in 1962 [4–7]. Initial devices were based on forward-biased GaAs pn junctions. Optical gain was provided by electron–hole recombination in the depletion region, and optical feedback was provided by polished facets perpendicular to the junction plane. This type of homojunction design meant that the carrier confinement of those lasers was poor, and the high laser threshold prohibited laser operation at room temperature. The concept of using wider bandgap material as one or both cladding layers to improve the laser carrier confinement and thus to reduce the leakage current was first proposed in 1963 [8]. Optical mode confinement was also expected to improve because a

larger refractive index of the center active layer would provide a waveguide effect. Seven years later, a continuous wave (CW) GaAs/AlGaAs double-heterojunction (DH) semiconductor laser operating at room temperature was demonstrated using a liquid-phase epitaxial (LPE) growth technique [9, 10]. Commercial application of edge-emitting semiconductor lasers has since become practical. Today, the worldwide semiconductor laser annual sales revenue has exceeded $400 million [11].

In the late 1970s, Iga et al. [12] proposed a semiconductor laser oscillating perpendicular to the device surface plane to overcome the difficulties facing edge-emitting semiconductor lasers that oscillate in parallel to the device surface plane. This type of laser is termed VCSEL, pronounced "VIXEL." VCSELs have demonstrated many advantages over edge-emitting semiconductor lasers. First, the monolithic fabrication process and wafer-scale probe testing as per the silicon semiconductor industry substantially reduces the manufacturing cost because only known good devices are kept for further packaging [13, 14]. Second, a densely packed two-dimensional (2-D) laser array can be fabricated because the device occupies no larger of an area than a commonly used electronic device [15]. This is very important for applications in optoelectronic integrated circuits. Third, the microcavity length allows inherently single longitudinal cavity mode operation due to its large mode spacing. Temperature-insensitive devices can therefore be fabricated with an offset between the wavelength of the cavity mode and the active gain peak [16, 17]. Finally, the device can be designed with a low numerical aperture and a circular output beam to match the optical mode of an optical fiber, thereby permitting efficient coupling without additional optics [14, 18].

A conventional edge-emitting semiconductor laser utilizes its cleaved facets as laser cavity reflectors because the length of the active layer is usually several hundred micrometers. The long active length provides enough optical gain to overcome the cavity reflector loss even though the reflectivity of the facets is only ~30%. In comparison, a VCSEL needs both of its surfaces to be highly reflective to reduce the cavity mirror loss because its active layer is less than 1 μm thick. The first VCSEL was demonstrated with GaInAsP/InP in 1979, which operated pulsed at 77°K with annealed Au at both sides as reflectors [12]. A room-temperature pulsed-operating VCSEL was demonstrated with a GaAs active region in 1984 [19]. Room-temperature CW-operating GaAs VCSELs were achieved by improving both the mirror reflectivity and the current confinement [20].

Currently, an output power of more than 100 mW has been obtained from an InGaAs/GaAs VCSEL with GaAs/AlAs monolithic diffractive Bragg reflectors (DBR) [21, 22]. VCSELs with lasing threshold of sub-100 μA [23, 24] or wall-plug efficiency of more than 50% have been reported with lateral oxidized–A1 confinement blocks [25, 26]. Room-temperature CW InGaAsP/InP VCSELs have met some difficulties primarily due to a low index difference between GaInAsP and InP, which causes difficulty in preparing a highly reflective monolithic DBR

[27]. Nevertheless, CW InGaAsP VCSELs at 1.5 μm have been reported using GaAs/AlAs DBR mirrors [28, 29].

Two-dimensional (2-D) arrayed VCSELs can find important applications in stacked planar optics, such as the simultaneous alignment of a tremendous number of optical components used in parallel multiplexing lightwave systems and parallel optical logic systems, free space optical interconnects, and so on. High-power lasers can also be made with phase-locked 2-D arrayed VCSELs. Some 2-D arrayed devices have been demonstrated [15, 30, 31], and efforts have been made in coherently coupling these arrayed lasers [32, 33].

5.4.1. Edge-Emitting Lasers

Two types of lasers have been extensively studied. They include near-infrared $Al_yGa_{1-y}As/Al_xGa_{1-x}As$ ($x > y \geq 0$) and long wavelength $In_xGa_{1-x}As_yP_{1-y}/InP$ DH edge-emitting semiconductor lasers. The epitaxial structures for both types of lasers are similar. They are usually *n–p–p* type, *n–i–p* type, or *n–n–p* type. LPE was the dominant epitaxial growth technique during the 1970s and early 1980s for high-quality semiconductor laser material growth, but it has gradually been taken over predominantly by metal organic chemical vapor deposition (MOCVD) techniques. Molecular beam epitaxy (MBE) is also a growth technique that has been used to demonstrate semiconductor lasers in research and development environments with a limited commercial success. A $GaAs/Al_xGa_{1-x}As$ DH laser is shown in Fig. 5.5 consisting of multiple compound semiconductor layers grown on a n^+-type GaAs substrate. The *p*-type active layer of the GaAs semiconductor laser, in which stimulated emission is amplified, is made of GaAs doped by Be, C, or Zn at $1 \times 10^{17}\,cm^{-3}$. The active layer thickness ranges from 50 to 2000 Å. On top of the active layer is a *p*-type $Al_xGa_{1-x}As$ cladding layer doped to $1 \times$

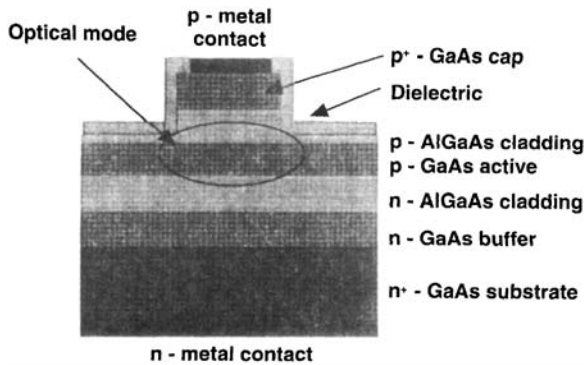


Figure 5.5 Schematic diagram of a ridge waveguide GaAs/AlGaAs double-heterojunction (DH) Laser.

$10^{18}\,cm^{-3}$ using Be, C, or Zn. Below the active layer is an *n*-type $Al_xGa_{1-x}As$ cladding layer doped to $1 \times 10^{18}\,cm^{-3}$ using Si, Sn, or Te. Cladding layer thickness usually ranges from 1500 Å to 1 μm. Within each $Al_xGa_{1-x}As$ layer, *x* is the value of aluminum mole fraction. When *x* is 0.3, for example, the energy bandgap of $Al_xGa_{1-x}As$ is approximately 1.8 eV–0.4 eV wider than that of GaAs active layer. When the pn junction is forward biased, electrons in the *n*-type cladding region are injected into the *p*-type active region. With a *p*-type semiconductor of wide bandgap on the other side of the pn junction, the injected minority carriers are mostly confined within the *p*-type active region. This carrier confinement allows population inversion to occur and optical gain to increase efficiently. In addition, the refractive index of GaAs is higher than that of $Al_xGa_{1-x}As$, and this acts like a waveguide to confine the majority of generated light within the GaAs active layer. The light that is not confined and penetrates into the $Al_xGa_{1-x}As$ cladding layers will not be absorbed by the cladding materials because of the wider bandgap and will therefore benefit laser action.

Although an epitaxy-ready GaAs substrate can be of good quality, there may exist some level of surface defects due to either grown-in crystal defects or mechanical polishing. Microscopic substrate surface flatness and native oxide on the substrate surface are also of great concern because they add a degree of difficulty to the epitaxial growth. To ensure a high-quality epitaxial crystal structure, the substrate usually first goes through the cleaning procedure, and a GaAs buffer layer with the same type of doping as that in the substrate is then deposited before the growth of any DH structure. There are also reports that high-quality epitaxial structures can be grown without buffer layers if substrates are thoroughly cleaned before the growth [34]. For the device structure shown in Fig. 5.5, the GaAs buffer layer is doped with a *n*-type dopant such as Si, Sn, or Te, typically having a concentration on the order of $1 \times 10^{18}\,cm^{-3}$. Its thickness ranges from 100 nm to several micrometers. The GaAs cap layer on top of the DH structure is very heavily doped with a *p*-type dopant such as Be, C, or Zn, typically at a concentration above $1 \times 10^{19}\,cm^{-3}$. This permits a low-resistivity metal Ohmic contact to be used for electrical conduction. Although a higher impurity doping concentration helps reduce the device series resistance, there is a limit at which the impurities can be incorporated into the host crystal structure. The adverse effect of overdoping is to form impurity clusters that behave as nonradiative recombination centers in the active region and the cladding layers, introducing internal optical loss or so-called free carrier absorption. The net effect is that the laser threshold current will increase. Therefore, the doping level at each layer of the device structure should be carefully adjusted to achieve the optimum designed laser performance.

The double-heterostructure semiconductor laser represents the single largest constituent of today's total semiconductor laser production because of its application in compact disk (CD) players and CD data storage. The current CD laser market volume is greater than 80 million units per year [11]. The high volume of

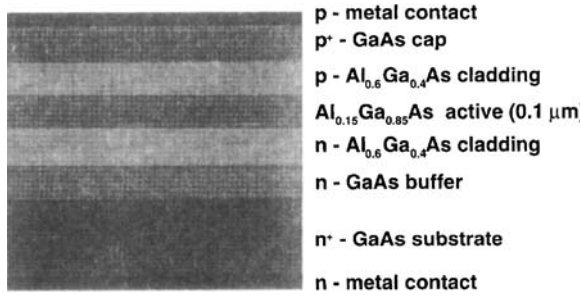

p - metal contact
p+ - GaAs cap
p - Al0.6Ga0.4As cladding
Al0.15Ga0.85As active (0.1 μm)
n - Al0.6Ga0.4As cladding
n - GaAs buffer
n+ - GaAs substrate
n - metal contact

Figure 5.6 Schematic diagram of an epitaxial structure for an edge-emitting semiconductor laser at 780 nm.

this market has driven the unit cost of a packaged laser to less than $1, with large-volume pricing of approximately $1. This has allowed businesses in the fiber optics market to use the low-cost CD laser in a historically high-cost environment. The CD laser has been used very successfully for some short-distance optical data links and has decreased laser wavelength specification from 850 nm for GaAs DH laser to 780 nm for CD laser [35, 36]. Because a CD laser operates at a wavelength of 780 nm, its active layer is made of AlGaAs with an A1 mole fraction of approximately 15%. For example, a CD laser may have an epitaxial structure consisting of 0.1-μm-thick undoped $Al_{0.15}Ga_{0.85}As$ active layer sandwiched between a *n*-type $Al_{0.6}Ga_{0.4}As$ cladding layer and a *p*-type $Al_{0.6}Ga_{0.4}As$ cladding layer grown on a *n*-type GaAs substrate, as shown in Fig. 5.6 [37]. The device fabrication is similar to any other type of DH lasers. Typically, a CD laser has a threshold of 20–50 mA at room temperature and is operated at an output power of approximately 3–5 mW, as shown in Fig. 5.7. The laser wavelength is usually between 770 and 795 nm. The low-threshold CD laser is used mostly for portable consumer electronics powered by batteries. When used in a CD player, a CD laser is usually designed to operate at multimode or self-pulsation mode in GHz range in order to reduce the feedback noise due to light reflection from a disk [37]. This type of CD laser, however, cannot be used for multi-Gb data communications.

When the active thickness of a DH laser is reduced to become comparable to the de Broglie wavelength [38], quantum mechanical effect starts to occur and the layer becomes a quantum well (QW). These QWs have been specifically used to design a new class of single quantum well (SQW) lasers. Two or more QWs can be placed between the two cladding layers to form a multiquantum well (MQW) laser. The layers separating the wells in the MQW laser are called barrier layers. Compared to a SQW laser, the MQW laser has a larger optical mode confinement factor, resulting in lower threshold carrier density and lower threshold

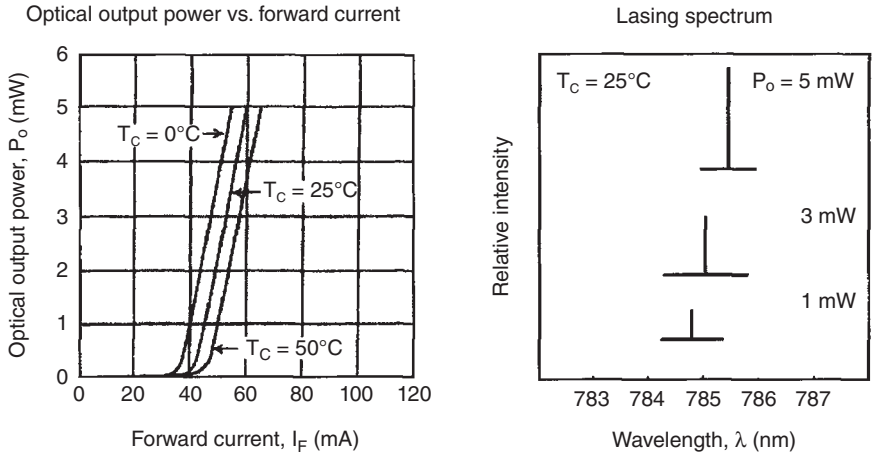


Figure 5.7 (Left) Output power vs. current of a CD laser, and (right) the correspondent laser spectrum.

current density. In comparison with a DH laser, the QW laser has a smaller active volume, a lower lasing threshold, and a higher differential gain, leading to increased relaxation oscillation frequency and reduced relative intensity noise. Quantum well lasers with small signal modulation frequencies above 20 GHz have been demonstrated [39, 40]. High-speed semiconductor lasers are important for large-bandwidth optical communications, as can be seen by the rapid deployment of optical fibers for transoceanic telecommunication cables and networking backbones throughout the Untied States and the world.

Shown in Fig. 5.8 is an energy band diagram of a GaAs/AlGaAs graded-dindex (GRIN) SQW laser structure [41]. The epitaxial layers are grown by MOCVD on a n^+-type GaAs substrate. The growth starts with a buffer layer of 2 μm, linear graded from GaAs to $Al_{0.6}Ga_{0.4}As$ and *n*-type doped at $2 \times 10^{18}\,cm^{-3}$. The *n*-type cladding layer of $Al_{0.6}Ga_{0.4}As$ has a thickness of 2 μm, doped at $2 \times 10^{18}\,cm^{-3}$. The *n*-side GRIN region is undoped with a thickness of 0.2 μm linear graded from $Al_{0.6}Ga_{0.4}As$ to $Al_{0.18}Ga_{0.82}AS$. The undoped GaAs QW is 50 Å thick. The *p*-side GRIN region is undoped with a thickness of 0.2 μm linear graded from $A1_{0.18}Ga_{0.82}As$ to $Al_{0.6}Ga_{0.4}As$. The *p*-type cladding layer of $Al_{0.6}Ga_{0.4}As$ has thickness of 2 μm, doped at $1 \times 10^{18}\,cm^{-3}$. The *p*-type GaAs cap layer is heavily doped at $1 \times 10^{19} cm^{-3}$, with a thickness of 0.5 μm. The GRIN layers in this structure are effective in improving both the carrier and the optical mode confinement. A threshold current density of only $200\,A/cm^2$ has been demonstrated from such a broad-area laser with a length of 400 μm and a stripe width of 180 μm [41].

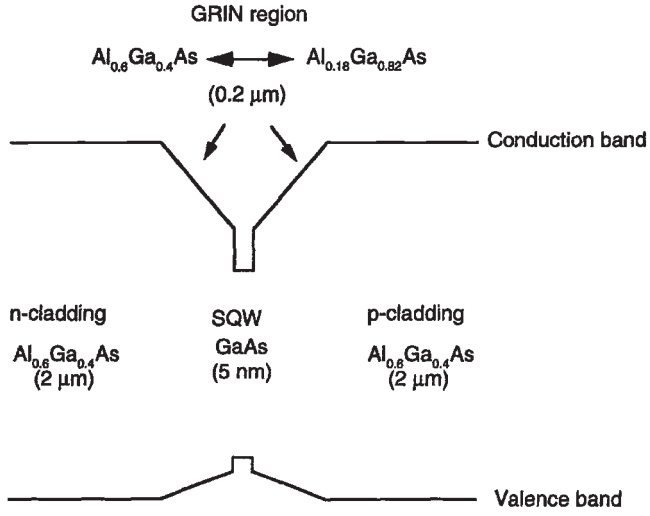


Figure 5.8 Energy band diagram of a graded-index (GRIN) single quantum well laser.

Further reduction of laser threshold current density can be achieved by using a strained QW active layer, such as $In_xGa_{1-x}As$ sandwiched between the GaAs barriers with $0 < x < 0.25$, or between the InP barriers with $x > 0.25$. The biaxial strain caused by the slight lattice mismatch between the two material systems alters the valence band edge by removing the degeneracy of the heavy hole and the light hole, resulting in reduced transparency carrier density and increased modal gain and thus reduced threshold current density [42]. The schematic energy band diagrams in *k* space for the strained gain medium are shown in Fig. 5.9. The strain in the active region also results in a larger differential gain, which helps improve the device operation bandwidth [40, 43]. Due to critical thickness constraints, the amount of strain in the active region and the active layer thickness, or the number of QWs with strained $In_xGa_{1-x}As$, is limited. To overcome the critical thickness barrier, strain-compensated active layers are used to increase the net total active thickness and the differential gain [44–47], thereby improving the optical mode confinement and reducing the laser threshold current density. The strain-compensated active structure consists of compressively strained quantum wells and tensile-strained barriers, or vice versa, so that the compressive strain and the tensile strain are mutually compensating for each other. Active layers exceeding the critical thickness can therefore be demonstrated without forming any lattice misfit dislocations.

In long-haul telecommunication systems, long-wavelength semiconductor lasers are of interest because of the minimum fiber dispersion at 1.3 μm and the

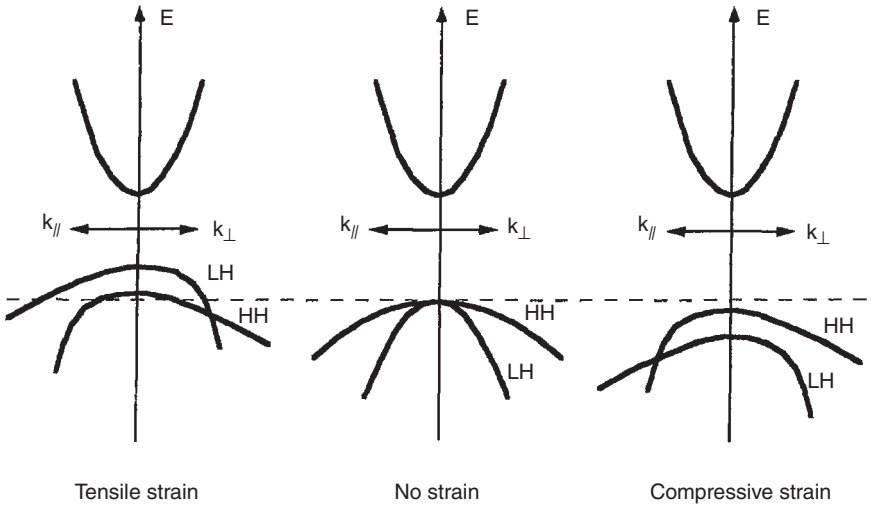


Figure 5.9 Schematic energy band diagram in *k* space showing the removal of degeneracy between heavy-hole (HH) valence band edge and light-hole valence band edge for both compressive- and tensile-strained InGaAs gain medium.

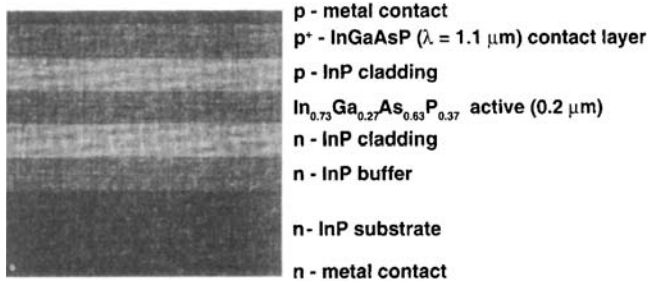

p - metal contact
p+ - InGaAsP (λ = 1.1 μm) contact layer
p - InP cladding
In0.73Ga0.27As0.63P0.37 active (0.2 μm)
n - InP cladding
n - InP buffer
n- InP substrate
n - metal contact

Figure 5.10 Schematic diagram of a 1.3-μm InGaAsP/InP DH laser epitaxial structure.

minimum fiber loss at 1.55 μm [48]. The dispersion-shifted fiber will have both the minimum dispersion and the minimum loss at 1.55 μm [49–51]. The long-wavelength semiconductor lasers are based on $In_xGa_{1-x}As_yP_{1-y}$ active layer lattice-matched to InP cladding layers [52, 53]. By varying the mole fractions x and y, any wavelength ranging between 1.1 and 1.6 μm can be selected. For example (Fig. 5.10), an InGaAsP/InP DH laser at 1.3 μm has an 0.2- to 0.3-μm-thick undoped $In_{0.73}Ga_{0.27}As_{0.63}P_{0.37}$ active layer, 3- or 4-μm *n*-type cladding layer doped by Sn at $2 \times 10^{18}\,cm^{-3}$, 2-μm *p*-type cladding layer doped by Zn at $1 \times 10^{18}\,cm^{-3}$, and 0.5-μm-thick *p*-type InGaAsP contact layer doped by Zn at $1 \times l0^{19}\,cm^{-3}$ [54].

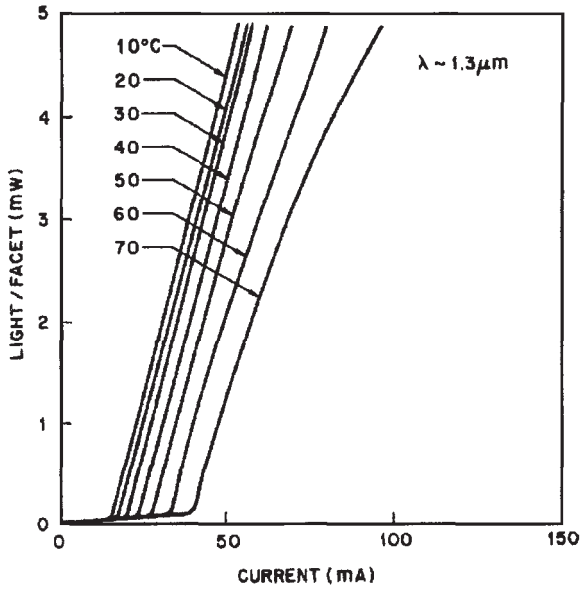


Figure 5.11 Laser output power vs. current of a 1.3-μm InGaAsP/InP DH.

The contact layer bandgap corresponds to a wavelength of 1.1 μm. An etched-mesa buried-heterostructure (EMBH) laser [54] made with this epitaxial structure has a threshold current of approximately 15 mA at room temperature and a single-mode output power of approximately 10 mW per facet, as shown in Fig. 5.11. One problem with the long-wavelength semiconductor lasers is the threshold current sensitivity to temperature, at room temperature (small T_0), due to poor carrier confinement and large Auger nonradiative recombination [55]. Improved thermal characteristics [56] and higher modulation speed [39] have been demonstrated when a MQW active layer is used for the long-wavelength semiconductor lasers.

Red visible semiconductor lasers operating in the range of 635–700 nm can find applications in bar-code scanners, laser printers, and laser pointers. They can also be used for plastic fiber data links because of the minimum loss at 650 nm in the plastic fiber [57]. With the emergence of digital video disk (DVD) technology for data storage [58], the market demand for both 635- and 650-nm semiconductor lasers is expected to soon catch up with the demand for the 780-nm CD lasers. Several material systems, such as AlGaAs [59], InGaAsP [60, 61], and InAlGaP [62–64], have been demonstrated to work in this wavelength region, but InAlGaP is regarded as the most appropriate material because it has a large direct energy bandgap while completely lattice-matched to a GaAs substrate. For

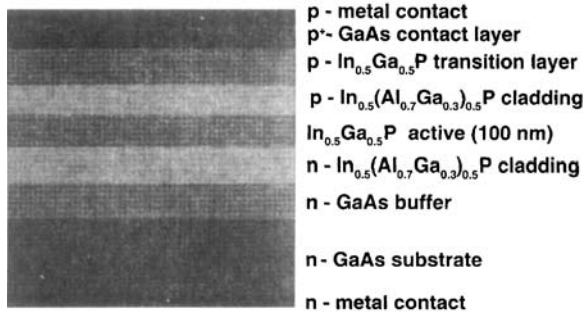

p - metal contact
p+- GaAs contact layer
p - In0.5Ga0.5P transition layer
p - In0.5(Al0.7Ga0.3)0.5P cladding
In0.5Ga0.5P active (100 nm)
n - In0.5(Al0.7Ga0.3)0.5P cladding
n - GaAs buffer
n - GaAs substrate
n - metal contact

Figure 5.12 Schematic diagram of a visible InGaP laser epitaxial structure.

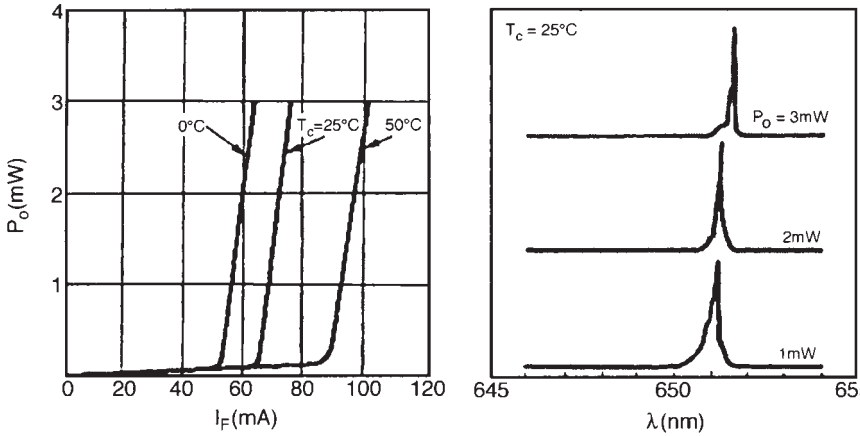


Figure 5.13 Laser output power vs. current of a 650-nm InGaP DH laser.

example (Fig. 5.12), an InAlGaP DH laser at 650 nm consists of an $In_{0.5}Ga_{0.5}P$ active layer sandwiched between the $In_{0.5}(Al_xGa_{1-x})0.5$ *p*-cladding layers grown on a GaAs substrate with $x = 0.7$. The active layer thickness is approximately 100 nm, and the cladding layer thickness is between 0.5 and 1 μm. Both the *n*-cladding and the p-cladding layers are doped at a concentration of $5 \times 10^{17}cm^{-3}$. Between the *p*-cladding layer and the *p*-GaAs contact layer is a *p*-type InGaP layer for improving the device I–V characteristics. This *p*-type InGaP intermediate layer is doped at $5 \times 10^{17}cm^{-3}$, and the *p*-type GaAs contact layer is doped as high as possible. The typical light output vs. current characteristics is shown in Fig. 5.13.

High-temperature performance and high-power operation have been the concerns for InAlGaP visible semiconductor lasers due to carrier leakage into the

p-cladding layer. Methods utilizing components, such as strained active layer [65, 66], off-angle substrate [67, 68], MQW active structure [69], and multiquantum barrier structure [70], have been developed to improve the laser performance.

The new DVD standard has increased the data storage capacity from 650 Mb to 4.7 Gb on a single-sided disk 12 cm in diameter. This storage capacity increase is attributed more to the tightening of system margin than to the shortening of laser wavelength from 780 to 635 or 650 nm. A 135-min high-definition motion picture, however, needs a storage capacity of 15 Gb. Because the current DVD standard has squeezed the system margin to the minimum, the future generation of DVD technology will rely to a great extent on laser wavelength shortening to expand the storage capacity in order to maintain the same disk size. Several groups have been investigating green/blue lasers using wide bandgap II–VI compound materials such as ZnCdSe/ZnSSe/ZnMgSSe grown on GaAs substrate [71–73], and good performance lasers have been demonstrated. The device, however, suffers serious reliability problems because of stacking fault-like defects that occur at and near the heterointerface between the GaAs substrate and the II–VI materials. Most II–VI semiconductor lasers degrade rapidly within minutes when running at CW. By reducing the grown-in defects, the device CW operation lifetime has been extended to 100 h [74] but is not yet long enough for any commercial applications.

Recent advancement in blue LED devices based on III–nitride materials [75] has prompted research in the blue/violet semiconductor lasers using InGaN/AlGaN MQW [76, 77] or InGaN/GaN DH structures [78]. The III–nitride epitaxial structures are grown on *c*-plane or *a*-plane sapphire substrates with a thick GaN buffer layer in between because no lattice-matched substrates are available. Lasers on a spinel ($MgA1_2O_4$) substrate have also been demonstrated [79]. Crystal quality, *p*-contact resistivity, carrier and current confinement, and facet mirror reflectivity have been the four major problems in the III–nitride semiconductor laser development [80]. Continuous wave operation at room temperature has been achieved at a wavelength at approximately 400 nm by improving the *p*-contact resistance and thus reducing the device operating voltage [81]. The device has a threshold current of approximately 3 or 4 kA/cm^2 and a lifetime of approximately 20 h at 1.5-mW constant power when running at room temperature. The improvement in reliability relies on further reducing the contact resistance and reducing the grow*n*-in crystal defects. The search for a lattice-matched substrate will help accelerate the device development cycles.

5.4.2. Vertical Cavity Surface-Emitting Lasers

The majority of the VCSELs being developed today are in the near-infrared wavelength range based on either GaAs/AlGaAs or strained InGaAs active materials. GaAs VCSELs at 850 nm are preferred as the light sources for short-

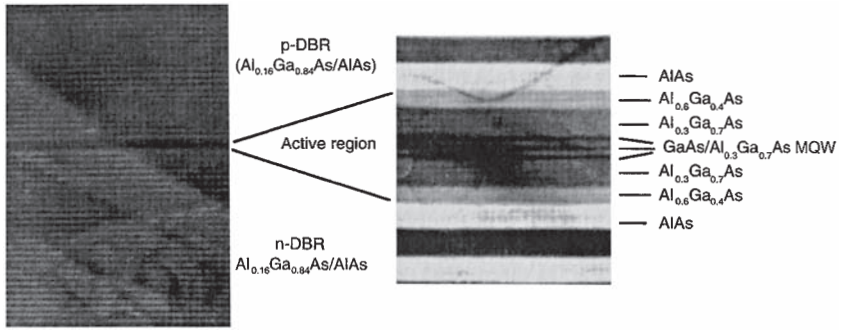


Figure 5.14 TEM photo of a GaAs VCSEL epitaxial layer structure.

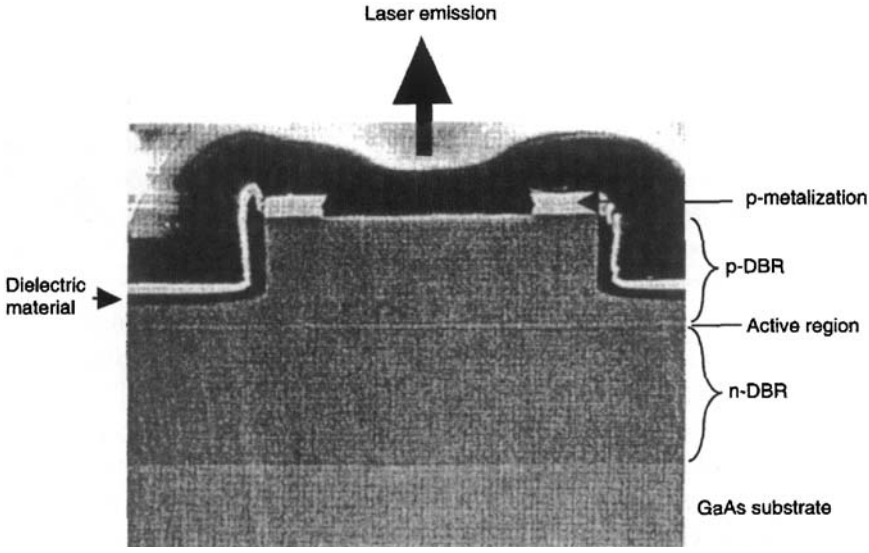


Figure 5.15 Cross-sectional SEM photo of an etched-mesa GaAs VCSEL structure.

distance optical communications because either Si or GaAs positive–intrinsic–negative (PIN) detectors can be used in the receiver end to reduce the total system cost. A typical GaAs VCSEL epitaxial layer structure is shown in Fig. 5.14, and the correspondent etched-mesa-type device structure is shown in Fig. 5.15. It includes three major portions: bottom Distributed Bragg Reflector (DBR), active region, and top DBR.

Generally, the epitaxial material is grown by either MOCVD or MBE techniques. The bottom DBR is *n*-type doped and grown on an *n*-type doped GaAs substrate. Typically, there is a GaAs buffer layer grown between the *n*-DBR and

the substrate. Silicon (Si) and selenium (Se) are two commonly used *n*-type dopants. The *n*-DBR is composed of 30.5 pair [82, 83] of $Al_{0.16}Ga_{0.84}As$/AlAs, which starts and stops with the AlAs layer alternated by the $Al_{0.16}Ga_{0.84}As$ layer. Each DBR layer has an optical thickness equivalent to a one-fourth of the designed lasing wavelength. The intrinsic cavity region is composed of two $Al_{0.3}Ga_{0.4}As$ spacer layers and three or four GaAs quantum wells, with each quantum well sandwiched between two $Al_{0.3}Ga_{0.7}As$ barriers. With the quantum well width of 100 Å and the quantum barrier width of 70 Å, the lasing wavelength is approximately 850 nm. The $Al_{0.6}Ga_{0.4}As$ spacer can be replaced by $Al_xGa_{1-x}As$ spacer, with x graded from 0.6 to 0.3 to form a graded-index separate confinement heterostructure. The total thickness of the spacers is such that the laser cavity length between the bottom and the top DBRs is exactly one wavelength or its multiple integer. The *p*-type doped DBR is grown on top of the active region. It consists of 22 pairs of $Al_{0.16}Ga_{0.84}As$/AlAs alternating layers that start with the AlAs layer and stop with the $Al_{0.16}Ga_{0.84}As$. Like the *n*-DBR, each layer has an optical thickness of one-fourth the wavelength. The most common *p*-type dopants utilized are carbon (C), zinc (Zn), and beryllium (Be). Typically, C is used for the *p* doping in the top DBR when using MOCVD growth techniques with the top several layers doped by Zn for better metalization contact [84]. Finally, a GaAs cap is used to prevent the top AlGaAs layer of the *p*-DBR from oxidization. The cap is highly *p* doped with Zn as the dopant and is kept to 100 Å thick because the material will be highly absorptive to the optical mode if it is too thick.

The function of a DBR in a VCSEL is equivalent to a cleaved facet in an edge-emitting laser: to reflect part of the laser emission back into the laser cavity and to transmit part of the laser emission as the output. It is in essence similar to a dielectric mirror. A pair of low-refractive index and high-refractive index layers with the optical thickness of each layer to be one-fourth of a specific wavelength will enhance the reflectivity to this wavelength. Many pairs of such alternating layers stacked together will form a mirror with its reflectivity reaching above 99% centered at the designed wavelength with a certain bandwidth. The mirror can be designed to achieve higher reflectivity and wider reflection bandwidth with a larger refractive index difference. $Al_xGa_{1-x}As$ is an ideal semiconductor material that can be monolithically grown on a GaAs substrate to provide a similar function as a dielectric mirror because the refractive index of this material can vary continuously from 3.6 to 2.9, with x varying from 0 to 1. For example, if one-fourth wavelength-thick layer is made of GaAs and the other one-fourth wavelength-thick layer is made of MAS, 20 pairs of such alternating layers will provide a reflectivity of 99% at the desired wavelength with a bandwidth of larger than 70 nm, as long as the desired wavelength is longer than the GaAs absorption band edge, which is approximately 875 nm at room temperature.

Although reflectivity of a natural cleaved facet in an edge-emitting laser is only approximately 30%, it is enough to ensure the lasing action due to the large

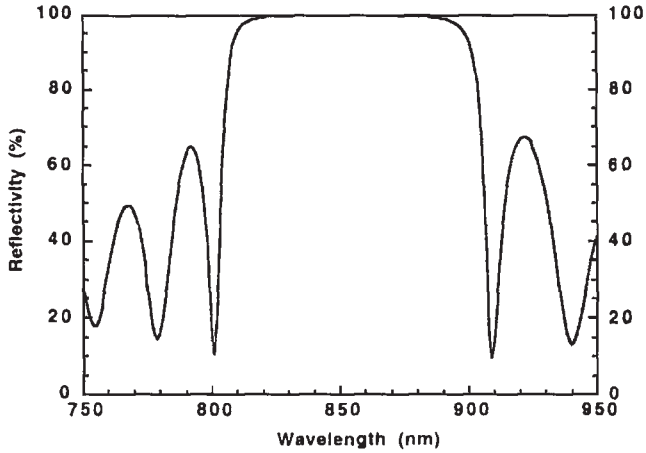


Figure 5.16 Reflectivity spectrum of 22 periods of $Al_{0.16}Ga_{0.84}As$/AlAs one-fourth wavelength DBR mirror stacks centered at a wavelength of 850 nm.

net gain provided by the long active length, which is usually several hundred micrometers. The reflectivity of a VCSEL DBR has to be more than 99% to reduce the cavity loss due to the extremely short gain length in the active region. For the example shown in Fig. 5.14, the optical gain is provided by the three GaAs quantum wells, with each quantum well thickness of only 100 Å. The net gain length of the VCSEL is thus only one-tenth of 1000 of that of an edge-emitting laser. Due to the reasonable refractive index difference between $Al_{0.16}Ga_{0.84}As$ and AlAs, the 22-pair DBR in this VCSEL has a reflectivity of 99.9% at 850 nm with a bandwidth of about 70 nm (Fig. 5.16). The large bandwidth provides some tolerance to any variation in the lasing wavelength due to growth variation in quantum well thickness and laser cavity length, the center wavelength mismatching between the top and the bottom DBRs, and the growth reactor/tooling variance. A typical GaAs VCSEL output power vs input current is shown in Fig. 5.17 for a mesa diameter of 10 μm and a laser emission aperture of 7 μm.

Although the structure shown in Fig. 5.14 is for GaAs VCSELs, many others have been working with strained InGaAs VCSELs operating at approximately 980 nm. The strain in the active region provides higher gain, which allows lower lasing threshold, and higher differential gain, which allows larger intrinsic modulation bandwidth. In addition, the InGaAs VCSEL has a wavelength transparent to the GaAs substrate, allowing the light emission toward the substrate side and thus the epitaxial side down packaging scheme. In this way, heat generated in the *p*-DBR mirror and the active junction region can be dissipated more efficiently, resulting in lower junction temperature and higher output power. A disadvantage

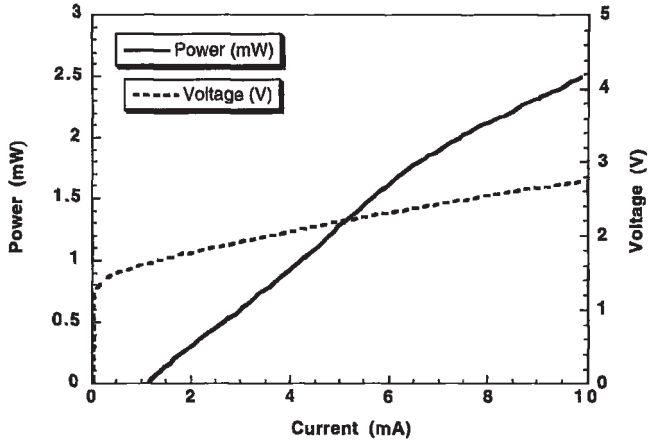


Figure 5.17 A GaAs VCSEL output power vs. input current for an etched-mesa structure with a mesa diameter of 10 μm and an emission aperture of 7 μm. The laser wavelength is approximately 850 nm.

of this type of VCSEL is that the low-cost Si or GaAs PIN detectors cannot be used when the VCSELs are used as the light sources for data communications. A typical InGaAs VCSEL structure consists of two DBR mirrors with reflection bands centered at approximately 970 nm [17, 85]. Each mirror is composed of one-quarter wavelength-thick layers alternating between AlAs and GaAs. The top *p*-doped DBR mirror contains 15 periods, whereas the bottom *n*-doped DBR mirror contains 18.5 periods. The cavity between the mirrors is filled by spacer layers of $Al_{0.3}Ga_{0.7}As$ that are used to center three 8-nm $In_{0.2}Ga_{0.8}As$ quantum wells (almost the thickness limit for coherently strained materials) separated by 10-nm GaAs barriers to form a one-wavelength-long cavity. The $Al_{0.3}Ga_{0.7}As$ spacer on the *p*-DBR mirror side is *p*-type doped, and the $Al_{0.3}Ga_{0.7}As$ spacer on the *n*-DBR mirror side is *n*-type doped. Above the *p*-DBR mirror stack, a heavily *p*-doped ($3 \times 10^{19}\,cm^{-3}$) GaAs phase-matching layer is deposited to provide a nonalloyed Ohmic contact to the hybrid Au mirror, which also acts as a *p* contact. The DBR mirrors are uniformly doped to $1 \times 10^{18}\,cm^{-3}$ except for the digital grading region, which is uniformly doped to $5 \times 10^{18}\,cm^{-3}$. The *n* dopant is Si and the *p* dopant is carbon. The whole epitaxial structure is grown by either MBE or MOCVD technique on a *n*-doped GaAs substrate. The device is designed to emit laser toward the substrate side.

Clearly, highly reflective semiconductor DBR mirrors are necessary to ensure low cavity loss, thus allowing the VCSEL to reach lasing threshold at a reasonable threshold carrier density level in the gain medium. Although the refractive index difference between the two constituents of the DBR structures is responsible for

high optical reflectivity, the accompanied energy bandgap difference that scales approximately linearly with the index difference results in electrical potential barriers in the heterointerfaces. These potential barriers impede the carrier flowing in the DBR structures and result in a large series resistance, especially in the *p*-type doping case. The large series resistance gives rise to thermal heating and thus deteriorates the laser performance.

The series resistance due to the heterojunctions in the DBR mirror can be minimized by grading and selectively doping the interfaces. In practice, the simplest approach to grade the interface is to introduce an extra intermediate-composition layer between the two alternating DBR constituents [86]. Instead of transiting directly from GaAs to AlAs, for example, a thin layer of $Al_{0.5}Ga_{0.5}As$ can be grown between the two to help smooth the interfaces. With the introduction of more $Al_xGa_{1-x}As$ layers of intermediate Al composition at each interface, greater performance over a single intermediate transition layer can be achieved. The advancement of MOCVD technology has allowed the continuous grading of an arbitrary composition profile. Very low-resistance DBRs have been achieved using this technique [87]. In the structure shown in Fig. 5.14, the heterointerfaces are linearly graded from 15% A1 composition to 100% over a distance of 12 nm. The total optical thickness of a pair of alternating layers is maintained to be one-half wavelength. The effect of this interface linear grading on the DBR mirror reflectivity is minimal. Likewise, heterointerface parabolic grading [88–90] and sinusoidal grading [9l] have also been used in some cases to flatten the valence band and therefore reduce the *p*-DBR series resistance.

The graded interfaces can sometimes be heavily doped ($5 \times 10^{l8}\,cm^{-3}$), whereas the remainder of the mirror is lightly doped ($1 \times 10^{l8}\,cm^{-3}$) to reduce scattering and free-carrier absorption loss in the mirror, and there is also a reduction in series resistance [92]. A simplified delta doping scheme [93, 94] has worked successfully to reduce the series resistance in the *p*-DBR mirror. In this scheme, *p* doping is carried out at interfaces where the nodes of the optical intensity are located, at levels as high as the crystal can incorporate. This heavy doping at the heterointerfaces causes the valence band edge to shift upward and the thermionic emission current to increase. The excess resistance at the higher bandgap side (AlAs) of the heterointerfaces is also reduced together with the relaxation of the carrier depletion in this region. Furthermore, the delta doping introduces a thinner potential barrier that allows for an increased tunneling current. Because the carrier density is increased only locally, the excess free-carrier absorption in the DBR mirror is minimized.

The intracavity metal contact technique [95–97] is an alternative approach to achieve low series resistance. This technique allows the electrical contact to bypass the resistive *p*-DBR mirror stack layers. Furthermore, the mirror stack above the metal contact does not require any doping, thereby reducing the intracavity free-carrier absorption loss and the optical scattering loss. Good performance

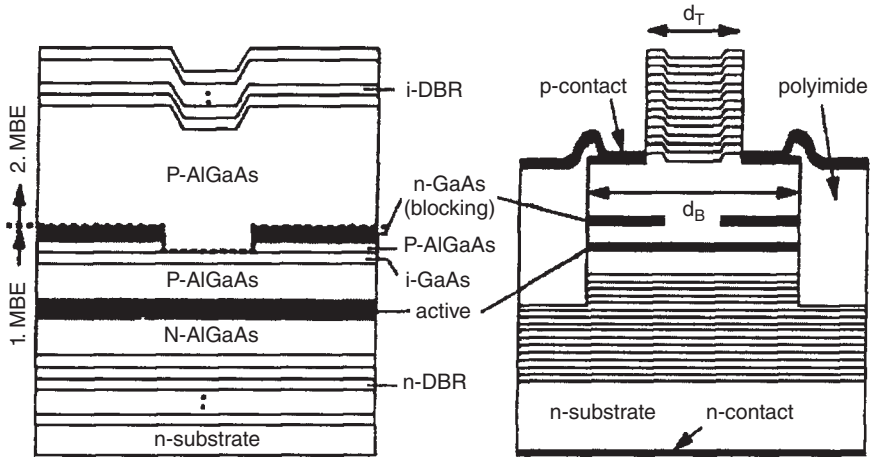


Figure 5.18 Schematic diagram of an intracavity contact VCSEL structure (after Ref. [96]).

VCSELs using this technique have been demonstrated in the research environment.

The intracavity metal contact structure starts with a conventional VCSEL design similar to that shown in Fig. 5.14 [95]. *p*-Type and *n*-type bulk layers that are multiples of half wavelengths are inserted on either side of the active region to provide electrical paths for the current to reach the active region from the ring contacts that are deposited on top of the inserted layers. A current blocking region must be formed to force the current into the optical mode. This current blocking region can be formed by ion implantation or undercutting using wet etching between the *p*-type insertion layer and the active region. A resistive layer is further introduced between the conductive current distribution layer and the active region to overcome any residual current crowding effects near the contact periphery so that the injected current can be more uniformly distributed in the optical mode area. The current blocking layer in the *p*-type region can also be a thin *n*-type GaAs or AlGaAs layer that is inserted into the top *p*-doped cladding layer [96]. A second growth is needed to complete the epitaxial structure after a current flow path is opened by either wet or dry etching. The final VCSEL device will have a reverse-biased pn junction inside the cavity in the *p*-doped region for current blocking (Fig. 5.18).

One of the simple approaches to make intracavity metal contact is to start with a VCSEL with an only partially grown *p*-DBR mirror stack (Fig. 5.19). After the metal contact has been deposited onto the *p*-DBR, a dielectric mirror stack consisting of quarter-wavelength-thick alternating layers TiO_2/SiO_2 is used to complete the device [97]. In this instance, resistance of $50\,\Omega$ has been achieved with a laser emission aperture of $5\,\mu m$ and a proton implantation aperture of $20\,\mu m$.

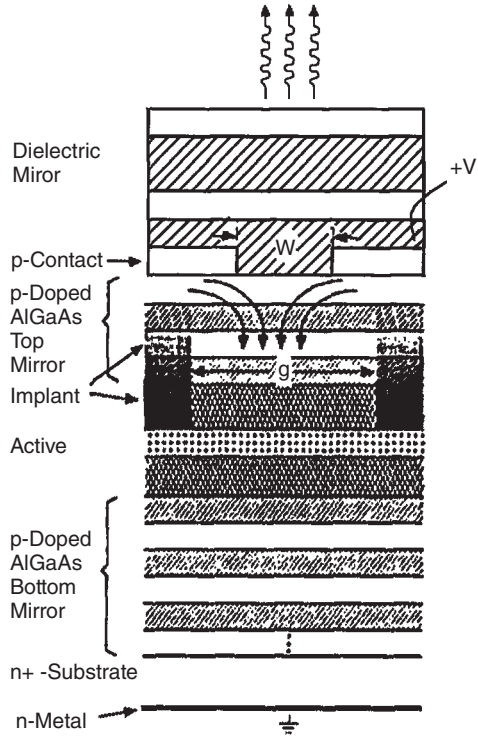


Figure 5.19 Schematic diagram of a VCSEL structure with partial monolithic semiconductor DBR and partial dielectric DBR on the *p*-doped contact side. (Reprinted with permission from Ref. [97]. Copyright 1995 American Institute of Physics.)

Within the microcavity structure of a VCSEL, only one Fabry–Perot cavity mode exists in the designed DBR reflective bandwidth. A laser can only be sustained at the wavelength of the cavity mode. Clearly, a temperature-insensitive VCSEL can be demonstrated by taking advantage of the microcavity mode characteristics [16]. Typically, the peak of the active gain profile shifts with the temperature at a rate of 3–5 Å/°C, and the cavity resonant mode shifts at a rate of 0.5–1 Å/°C [98]. If the resonant cavity mode is designed to initially sit at the longer wavelength side of the gain profile, the gain peak will gradually walk into the cavity mode with the rise of temperature (Fig. 5.20). Conversely, the gain peak usually decreases with the temperature. Together, the actual gain for the VCSEL cavity mode will vary little with temperature, and the VCSEL threshold current will stay almost constant within a certain temperature range. This temperature insensitivity allows the VCSEL to be designed to operate optimally at the system temperature midpoint region. For example, with modest speed optical

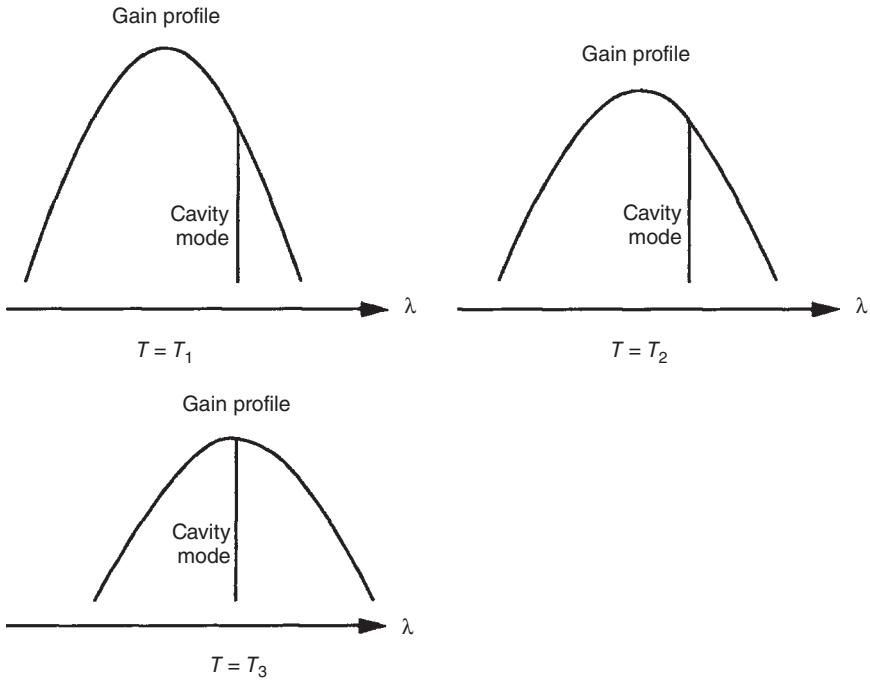


Figure 5.20 SEL cavity mode vs gain profile for temperatures T_1, T_2, and T_3, with $T_3 > T_2 > T_1$.

data links, the VCSELs can be designed to operate with minimum threshold current at approximately 40°C in a required working range of 0–70°C (Fig. 5.21) [17, 99]. The system can be implemented without any auto power control (APC) circuitry, thereby simplifying the packaging and reducing the system cost [14]. The application of this method has also allowed the demonstration of VCSELs operating at a record high temperature of 200°C [83].

Apart from VCSELs at 830–870 nm based on GaAs MQWs and VCSELs at 940–980 nm based on strained InGaAs MQWs, VCSELs operating at other wavelengths, such as 780 nm based on AlGaAs MQWs, 650–690 nm based on InAlGaP MQWs, and 1.3–1.5 nm VCSELs based on InGaAsP MQWs, have received attention in the research community.

The vast majority of the semiconductor laser market is at 780 nm, which is predominantly used for CD data storage and laser printing. As a result, the development of VCSELs at 780 nm is of strategic importance from a commercial standpoint. A typical VCSEL at 780 nm has an epitaxial layer structure similar to that of a VCSEL at 850 nm [100–102]. The larger bandgap requirement for 780 nm drives the MQW active region to the AlGaAs ternary system. The active region

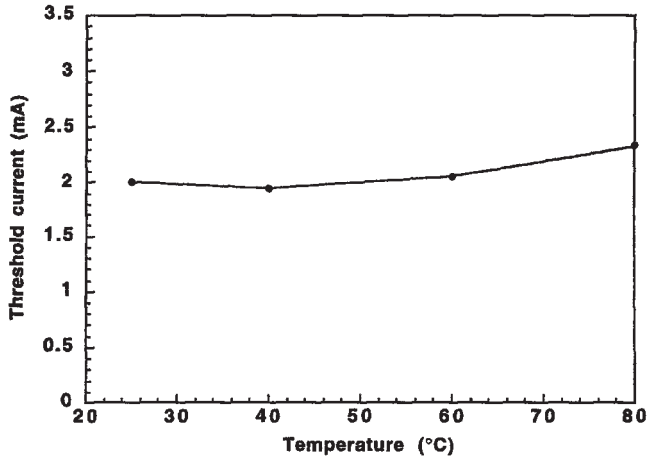


Figure 5.21 Threshold current of a typical GaAs VCSEL varying with ambient temperature with minimum threshold current at 40°C.

usually consists of three or four periods of $A1_{0.12}Ga_{0.88}As$ quantum wells sandwiched between the $Al_{0.3}Ga_{0.7}As$ barriers. The DBR mirror stack consists of 27 pairs of *p*-type doped $Al_{0.25}Ga_{0.75}As$/AlAs and 40 pairs of *n*-type doped $Al_{0.25}Ga_{0.75}As$/AlAs, with the bandwidth centered at 780 nm. The laser performance of a 780-nm VCSEL is similar to that of an 850-nm GaAs VCSEL (Fig. 5.22). The increased aluminum concentration in both the active region and the DBR mirror stack over that used in the 850-nm VCSEL raises a concern with the 780-nm VCSEL device reliability because of the poor edge-emitting semiconductor laser performance at 780 nm. No reliability data have been published so far for the 780-nm VCSELs, and study is ongoing to address the issue.

Red visible VCSELs are of interest because of their potential applications in plastic fiber, bar-code scanner, pointer, and most recently the DVD format optical data storage. The epitaxial structure of a red visible VCSEL is grown on a GaAs substrate misoriented 6° off (100) plane toward the nearest |111 > A or on a (311) GaAs substrate [103–106]. It consists of three or four periods of $In_{0.56}Ga_{0.44}P$QWs with InAlGaP or InAlP as barriers, InAlP as both *p*-type and *n*-type cladding layers, and two DBR mirrors (Fig. 5.23). The active QW layer is either tensile or compressive strained to enhance the optical gain. Typically, the QW thickness is 60–80 Å and the barrier thickness is 60–100 Å. The total optical cavity length including the active region and the cladding layers ranges from one wavelength or its multiple integer up to eight wavelengths. The DBR mirrors are composed of either InAlGaP/InAlP or $Al_{0.5}Ga_{0.5}As$/AlAs. The $Al_{0.5}Ga_{0.5}As$/AlAs DBR mirror performs better because of a relatively larger index difference between the two

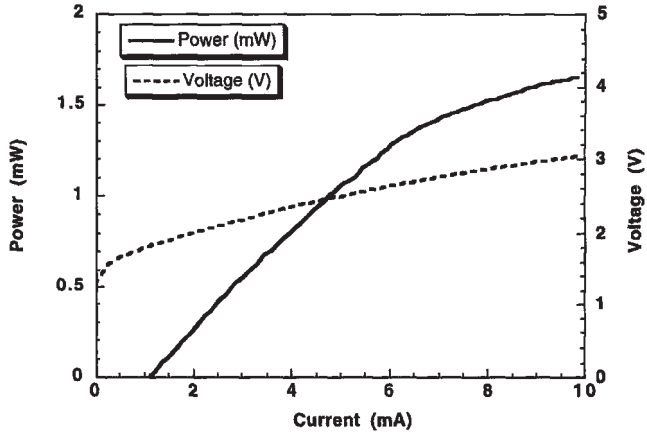


Figure 5.22 Etched-mesa structure VCSEL output power vs input current at a wavelength of 780 nm.

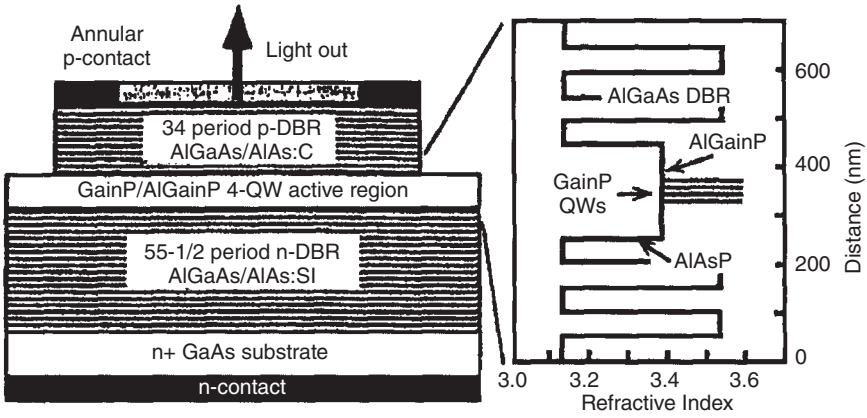


Figure 5.23 A visible VCSEL structure. (Reprinted with permission from Ref. [105]. Copyright 1995 American Institute of Physics.)

DBR constituents—thus a higher reflectivity and a wider bandwidth. In general, because the index difference between $Al_{0.5}Ga_{0.5}As$ and AlAs is much smaller than that used for the 850-nm VCSELs, more mirror pairs are needed to achieve the required DBR reflectivity. Typically, 55 pairs are needed for the *n*-DBR and 40 pairs are needed for the *p*-DBR to ensure a reasonable VCSEL performance. As a rule of thumb, the more pairs in the DBR mirror, the higher series resistance and thus more heat generated in the active region. This implies that the active

junction temperature will be higher. Currently, sub-mA threshold red VCSELs have been demonstrated. More than 5-mW output power from a red VCSEL has also been reported. Unfortunately, the carrier confinement of the red visible VCSELs is poor because of the smaller bandgap offset between the quantum well and the barrier and between the active and the cladding. Therefore, the red visible VCSELs are extremely temperature sensitive, and more studies are needed to improve the red visible VCSEL high-temperature performances. VCSELs with wavelengths shorter than 650 nm pose more problems because of even worse carrier confinements. Designing a VCSEL that can effectively confine the carriers in the active region is a challenging topic for today's research community.

Long-wavelength VCSELs at 1.3 and 1.55 μm have drawn attention because of their potential applications in telecommunications and medium- to long-distance data links, such as local area networks and wide area networks, where single-mode characteristics are required. The long-wavelength VCSELs are based on an InP substrate, with InGaAsP MQWs used as the active region. However, the lattice-matched monolithic InGaAsPDnP DBR mirrors do not have sufficient reflectivity for the long-wavelength VCSELs because of the small index difference between the two DBR mirror pair constituents, InGaAsP and InP. In addition, the Auger recombination-induced loss becomes evident due to smaller energy bandgap for the long-wavelength VCSELs. To overcome the difficulty, dielectric mirrors with 8.5 pairs of MgO/Si multilayers and Au/Ni/Au on the *p* side and 6 pairs of SiO_2/Si on the *n* side have been used instead of the semiconductor DBR. A continuous-wave 1.3-μm VCSEL has therefore been demonstrated at 14°C [107]. To further improve the device performance, wafer-fusing techniques have been adopted to bond GaAs/AlAs DBR mirrors onto a structure with an InGaAsP MQW active layer sandwiched between the InP cladding layers that are epitaxially grown on the InP substrate [108, 109]. The InP substrate is removed to allow the GaAs/AlAs DBRs to be bonded onto one or both sides of the InGaAsP active region (Fig. 5.24). Because the DBR mirrors are either *n*-type or *p*-type doped, the completed fused wafer can be processed like a regular GaAs VCSEL wafer. In this way, a 1.5-μm VCSEL has been successfully fabricated that operates CW up to 64°C [28, 29]. Manufacturing yield and reliability are still currently unknown with the VCSEL wafer fusion technique. For commercial interest, the CW operation must be driven to at least the 100°C range for the junction, in addition to a number of other issues such as wall-plug efficiency, reliability, and consistency.

Angle Polished Connector (APC) is one of the important features that is easily accomplished with edge-emitting lasers because of the backward emission that can be monitored from the cleaved facet. With VCSELs of wavelength shorter than 870 nm, the laser beam emits only toward the top epitaxy side. The backward emission is absorbed by the GaAs substrate, unless the substrate is removed. However, due to the unique vertical stacking feature of VCSELs, a detector can

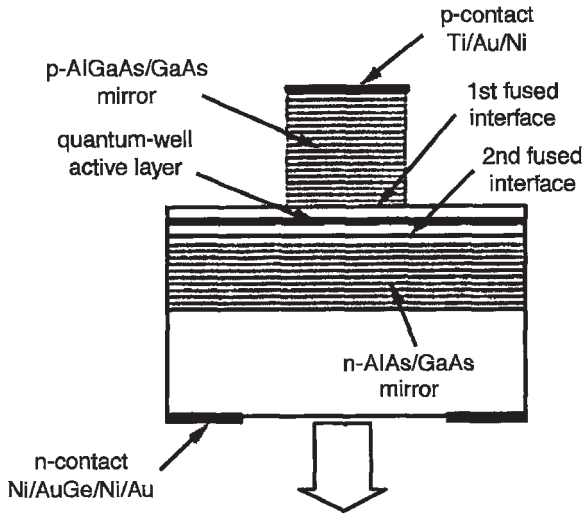


Figure 5.24 Schematic diagram of a wafer-fused long-wavelength VCSEL (after Ref. [28]).

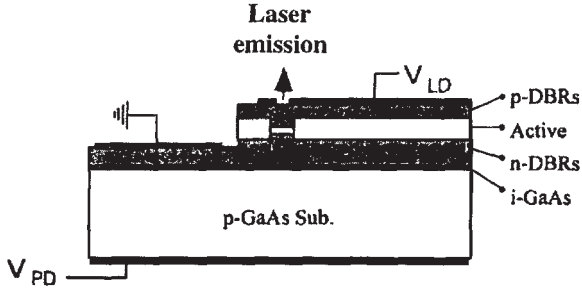


Figure 5.25 Schematic diagram of a VCSEL with integrated detector (after Ref. [102]).

be integrated underneath or above the VCSEL structure during the epitaxial growth [102, 110–113] (Fig. 5.25). For example, a VCSEL can start with a *p*-type GaAs substrate, with a PIN detector structure grown first on top of the substrate. The PIN detector has a GaAs intrinsic layer of approximately 1 μm and *p*-doped AlGaAs cladding of approximately 2000 Å between the substrate and the intrinsic absorption layer. The detector structure stops at a *n*-type doped cladding layer of approximately 2000 Å. A regular GaAs VCSEL epitaxial structure follows the PIN detector, with layers of *n*-DBR, *n* cladding, active, *p* cladding, and *p*-DBR grown in order. The detector cathode in this structure shares a common contact with the VCSEL cathode, with two independent anodes for both the PIN detector

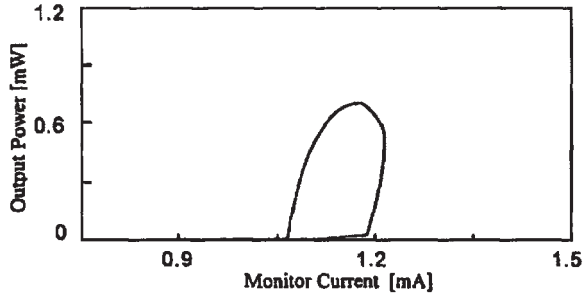


Figure 5.26 SEL output power in relationship with current response of an integrated detector (after Ref. [102]).

and the VCSEL. In practical applications, the anode of the detector can be either reverse biased or without any bias if detector speed is not a major concern. The VCSEL backward emission transmitted through the *n*-DBR is normally in proportion to the VCSEL forward emission. It will be received by the integrated PIN detector and generate a current. The VCSEL output power and the integrated PIN detector response are shown in Fig. 5.26. There is a one-to-one relationship between the PIN detector current and the VCSEL output power up to a certain point when the VCSEL output power saturates, but the detector current keeps rising due to the effect of spontaneous emission. Consequently, VCSEL operation with APC can be accomplished by monitoring the current variation generated in this detector when the VCSEL operates below the saturation [102, 110].

Super-low-threshold microcavity-type VCSELs have been proposed that utilize the spontaneous emission enhancement due to more spontaneous emission being coupled into the lasing mode [114, 115]. Although a thresholdless laser is theoretically possible when the spontaneous emission coupling effciency β is made approaching unity, the proposed structures are difficult to make in practice. One of the successful examples in research today is the use of oxidized lateral carrier confinement blocks by oxidizing an AlAs layer in the DBR or the cladding regions [23, 24, 116] (Fig. 5.27). This technology will be discussed in more detail in Chapter 8. Typically, sub-100-μA threshold can be achieved with this technique. A VCSEL with an extremely low threshold of 8.7 μA has been reported with an active area of $3\,\mu m^2$ [24]. It should be noted that there is still a debate on the exact mechanism that has generated this result. VCSELs with oxidized mirrors have been demonstrated with extremely simple epitaxy layers [117, 118]. In this structure, only four to six pairs of GaAs/AlAs DBR stacks are grown on one or both sides of an active region that is made of strained InGaAs MQWs at 970 nm (Fig. 5.28). The AlAs layers in the DBR mirrors are oxidized during the fabrication procedure. The extremely large index difference between GaAs and the

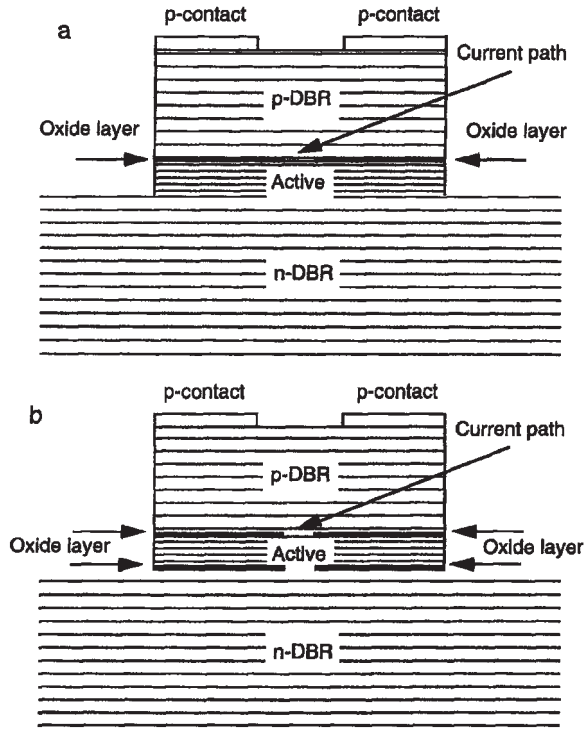


Figure 5.27 SEL with native aluminum oxide for lateral current confinement. (a) Current confinement on *p* side, and (b) current confinement on both *p* side and *n* side.

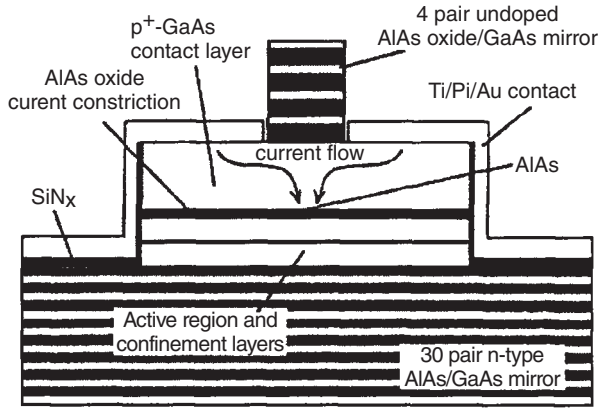


Figure 5.28 SEL with an AlAs oxide–GaAs DBR mirror (after Ref. [117]).

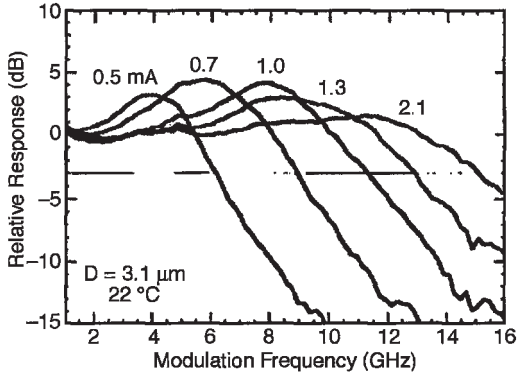


Figure 5.29 Small signal modulation response of a 3-μm VCSEL at various bias current. The maximum 3-dB bandwidth is approximately 15 GHz (after Ref. [120]).

oxidized AlAs layer makes it possible that only four pairs of GaAs/AlAs stacks will provide sufficiently high reflectivity with very large bandwidth for proper device operation. The VCSEL electrical contacts in this case will have to be made laterally inside the cavity as opposed to those at the top of the DBR mirror stacks because electrical conduction through the DBR mirror is prohibited once the AlAs constituent of the mirror is oxidized.

High-speed data transmission requires that a VCSEL be modulated at multi-GHz. The cavity volume of a VCSEL is significantly smaller than that of an edge-emitting laser, resulting in a higher photon density in the VCSEL cavity. The resonance frequency of a semiconductor laser typically scales as the square root of the photon density, thus indicating that a VCSEL has a potential advantage in high-speed operation. However, the parasitic series resistance caused by the semiconductor DBR and the device heating limit the maximum achievable VCSEL modulation bandwidth. Currently, a modulation speed of larger than 16 GHz has been reported with an oxideconfined VCSEL at a current of 4.5 mA [119]. Modeling results indicate that a gain compression limited-oxide VCSEL with a diameter of 3 μm has an intrinsic 3-dB bandwidth of 45 GHz [120] and a measured 3-dB bandwidth of 15 GHz at 2.1 mA due to the parasitic resistance and the device heating (Fig. 5.29).

REFERENCES

1. Gowar, J. 1984. *Optical communication systems.* Englewood Cliffs, N.J.: Prentice Hall.
2. Miller, S. E., and A. G. Chynoweth, eds. 1979. *Optical fiber telecommunications.* New York: Academic Press.
3. Lasky, R., U. Osterberg, and D. Stigliani, eds. 1995. *Optoelectronics for data communication.* New York: Academic Press.

4. Hall, R. N., G. E. Fenner, J. D. Kingsley, T. J. Soltys, and R. 0. Carlson. 1962. Coherent light emission from GaAs junctions. *Phys. Rev. Lett.* 9:366.
5. Nathan, M. I., W. P. Dumke, G. Burns, F. H. Dill, Jr., and G. Lasher. 1962. Stimulated emission of radiation from GaAs pn junctions. *Appl. Phys. Lett.* 1:62.
6. Holonyak, N., Jr., and S. F. Bevacqua. 1962. Coherent (visible) light emission from Ga $(Al_{1-x}P_x)As$ junctions. *Appl. Phys. Lett.* 1:82.
7. Quist, T. M., R. H. Rediker, R. J. Keyes, W. E. Krag, B. Lax, A. L. McWhorter, and J. J. Zeiger. 1962. Semiconductor maser of GaAs. *Appl. Phys. Lett.* 1:91.
8. Kroemer, H. 1963. A proposed class of heterojunction injection lasers. *Proc. IEEE* 51:1782.
9. Hayashi, I., M. B. Panish, P. W. Foy, and S. Sumuski. 1970. Junction lasers which operate continuously at room temperature. *Appl. Phys. Lett.* 17:109.
10. Alferov, Zh. I., V. M. Andreev, D. Z. Garbuzov, Yu. V. Zhilyaev, E. P. Morozov, E. L. Portnoi, and V. G. Tiiofim. 1971. Investigation of the influence of the AlAs-GaAs heterostructure parameters on the laser threshold current and the realization of continuous emission at room temperature. *Sov. Phys. Semiconductor* 4:1573.
11. Anderson, S. G. 1996. Annual review of laser markets. *Laser Focus World* 32:50.
12. Soda, H., K. Iga, C. Kitahara, and Y. Suematsu. 1979. GaInAsP/InP surface emitting injection lasers. *Jpn. J. Appl. Phys.* 18:2329.
13. Iga, K., F. Koyama, and S. Kinoshita. 1988. Surface emitting semiconductor lasers. *IEEE J. Quantum Electron.* QE-24:1845.
14. Lebby, M., C. A. Gaw, W. B. Jiang, P. A. Kiely, C. L. Shieh, P. R. Claisse, J. Ramdani, D. H. Hartman, D. B. Schwartz, and J. Grula. 1996. Use of VCSEL arrays for parallel optical interconnects. *Proc. SPIE* 2683:81.
15. Orenstein, M., A. C. Von Lehmen, C. Chang-Hasnain, N. G. Stoffel, J. P. Harbison, and L. T. Florez. 1991. Matrix addressable vertical cavity surface emitting laser array. *Electron. Lett.* 27:437.
16. Shieh, C. L., D. E. Ackley, and H. C. Lee. 1993. Temperature insensitive vertical cavity surface emitting laser. U.S. Patent No. 5,274,655.
17. Young, D. B., J. W. Scott., F. H. Peters, B. J. Thibeault, S. W. Corzine, M. G. Peters, S. L. Lee, and L. A. Coldren. 1993. High-power temperature insensitive gain-offset InGaAs/GaAs vertical-cavity surface-emitting lasers. *IEEE Photon. Tech. Lett.* 5:129.
18. Tai, K., G. Hasnain, J. D. Wynn, R. J. Fischer, Y. H. Wang, B. Weir, J. Gamelin, and A. Y. Cho. 1990. 90% coupling of top surface emitting GaAs/AlGaAs quantum well laser output into 8 μm diameter core silica fibre. *Electron. Lett.* 26:1628.
19. Iga, K., S. Ishikawa, S. Ohkouchi, and T. Nishimura. 1984. Room temperature pulsed oscillation of GaAlAs/GaAs surface emitting laser. *Appl. Phys. Lett.* 45:348.
20. Koyama, F. S. Kinoshita, and K. Iga. 1988. Room-temperature CW operation of GaAs vertical cavity surface emitting laser. *Trans. Inst. Electron. Commun. Eng. Jpn.* E71:1089.
21. Peters, E H., M. G. Peters, D. B. Young, J. W. Scott, B. J. Thibeault, S. W. Corzine, and L. A. Coldren. 1993. High power vertical cavity surface emitting lasers. *Electron. Lett.* 29:200.
22. Grabherr, M., B. Weigl, G. Reiner, R. Michalzik, M. Miller, and K. J. Ebeling. 1996. High power top-surface emitting oxide confined vertical-cavity laser diodes. *Electron. Lett.* 32:1723.
23. Huffaker, D. L., J. Shin, and D. G. Deppe. 1994. Low threshold halfwave vertical-cavity lasers. *Electron. Lett.* 30:1946.
24. Yang, G. M., M. H. MacDougal, and P. D. Dapkus. 1995. Ultralow threshold current vertical-cavity surface-emitting lasers obtained with selective oxidation. *Electron. Lett.* 31:886.
25. Lear, K. L., K. D. Choquette, R. P. Schneider, S. P. Kilcoyne, and K. M. Geib. 1995. Selectively oxidised vertical cavity surface emitting lasers with 50% power conversion efficiency. *Electron. Lett.* 31:208.
26. Jäger, R., M. Grabherr, C. Jung, R. Michalzik, R. Reiner, B. Weigl, and K. J. Ebeling. 1997. 57% wallplug efficiency oxide-confined 850 nm wavelength GaAs VCSELs. *Electron. Lett.* 33:4.

27. Iga, K. 1992. Surface emitting lasers. *Opt. Quantum Electron.* 24:S97.
28. Babic, D. I., K. Streubel, R. P. Mirin, N. M. Margalit, J. E. Bowers, E. L. Hu, D. E. Mars, L. Yang, and K. Carey. 1995. Room-temperature continuouswave operation of 1.54-μm vertical-cavity lasers. *IEEE Photon. Tech. Lett.* 7:1225.
29. Margalit, N. M., D. I. Babic, K. Streubel, R. P. Mirin, R. L. Naone, J. E. Bowers, and E. L. Hu. 1996. Submilliamp long wavelength vertical cavity lasers. *Electron. Lett.* 32:1675.
30. Vakhshoori, D., J. D. Wynn, G. J. Zydik, and R. E. Leibenguth. 1993. 8×18 top emitting independently addressable surface emitting laser arrays with uniform threshold current and low threshold voltage. *Appl. Phys. Lett.* 62:1718.
31. Uchiyama, S., and K. Iga. 1985. Two-dimensional array of GaInAsP/InP surface-emitting lasers. *Electron. Lett.* 21:162.
32. Deppe, D. G., J. P. van der Ziel, N. Chand, G. J. Zydzik, and S. N. G. Chu. 1990. Phase-coupled two-dimensional $Al_xGa_{1-x}As$-GaAs vertical-cavity surface-emitting laser array. *Appl. Phys. Lett.* 56:2089.
33. Orenstein, M., E. Kapon, N. G. Stoffel, J. P. Harbison, L. T. Florez, and J. Wullert. 1991. Two-dimensional phase-locked arrays of vertical-cavity semiconductor lasers by mirror reflectivity modulation. *Appl. Phys. Lett.* 58:804.
34. Iizuka, K., K. Matsumaru, T. Suzuki, H. Hirose, K. Suzuki, and H. Okamoto. 1995. Arsenic-free GaAs substrate preparation and direct growth of GaAs/AlGaAs multiple quantum well without buffer layers. *J. Cryst. Growth* 150:13.
35. Cheng, W. H., and J. H. Bechtel. 1993. High-speed fibre optic links using 780 nm compact disc lasers. *Electron. Lett.* 29:2055.
36. Soderstrom, R. L., S. J. Baumgmer, B. L. Beukema, T. R. Block, and D. L. Karst. 1993. CD lasers optical data links for workstations and midrange computers. *ECTC'93,* 505, June, Orlando.
37. Nakata, N. 1987. Laser diodes have low noise and low astigmatism. *JEE,* August, 49.
38. Wang, S. 1989. *Fundamentals of semiconductor theory and device physics,* 51. Englewood Cliffs, N.J.: Prentice Hall.
39. Morton, P. A., R. A. Logan, T. Tanbunek, P. F. Sciortino, A. M. Sergent, R. K. Montgomery, and B. T. Lee. 1993. 25 GHz bandwidth 1.55-μm GaInAsP p-doped strained multiquantum-well lasers. *Electron. Lett.* 29:136.
40. Ralston, J. D., E. C. Larkins, K. Eisele, S. Weisser, S. Buerkner, A. Schoenfelder, J. Daleiden, K. Czotscher, I. Esquivias, J. Fleissner, R. E. Sah, M. Maier, W. Benz, and J. Rosenzweig. 1996. Advanced epitaxial growth and device processing techniques for ultrahigh-speed (>40 GHz) directly modulated semiconductor lasers. *Proc. SPIE* 2683:30.
41. Hersee, S. D., B. de Cremoux, and J. P. Duchemin. 1984. Some characteristics of the GaAs/GaAlAs graded-index separate-confinement heterostructure quantum well laser structure. *Appl. Phys. Lett.* 44:476.
42. Coleman, J. J. 1995. Quantum-well heterostructure lasers. In *Semiconductor lasers: Past, present, and future,* ed. G. P. Agrawal, Chapter 1. Woodbury, N.Y.: AIP Press.
43. Nagarajan, R., T. Fukushima, J. E. Bowers, R. S. Geels, and L. A. Colden. 1991. High-speed InGaAs/GaAs strained multiple quantum well lasers with low damping. *Appl. Phys. Lett.* 58:2326.
44. Miller, B. I., U. Koren, M. G. Young, and M. D. Chien. 1991. Strain-compensated strained-layer superlattices for 1.5 μm wavelength lasers. *Appl. Phys. Lett.* 58:1952.
45. Zhang, G., and A. Ovtchinnikov. 1993. Strain-compensated InGaAs/GaAsP/GaInAsP/GaInP quantum well lasers (1 ~ 0.98 μm) grown by gas-source molecular beam epitaxy. *Appl. Phys. Lett.* 62:1644.
46. Tsuchiya, T., M. Komori, R. Tsuneta, and H. Kakibayashi. 1994. Investigation of effect of strain-compensated structure and compensation limit in strained-layer multiple quantum wells. *J. Cryst. Growth* 145:371.

47. Bessho, Y., T. Uetani, R. Hiroyama, K. Komeda, M. Shono, A. Ibaraki, K. Yodoshi, and T. Niina. 1996. Self-pulsating 630 nm band strain-compensated MQW AlGaInP laser diodes. *Electron. Lett.* 32:667.
48. Nagel, S. R., J. B. MacChesney, and K. L. Walker. 1985. Modified chemical vapor deposition. In *Optical fiber communications*, ed. T. Y. Li, Vol. 1, Chap. 1. Orlando, FL.: Academic Press.
49. Cohen, L. G., C. Lin, and W. G. French. 1979. Tailoring zero chromatic dispersion into the 1.5–1.6 μm low-loss spectral region of single-mode fibres. *Electron. Lett.* 15:334.
50. Tsuchiya, H., and N. Imoto. 1979. Dispersion-free single-mode fibre in 1.5 μm wavelength region. *Electron. Lett.* 15:476.
51. Okamoto, K., T. Edahiro, A. Kawana, and T. Miya. 1979. Dispersion minimization in single-mode fibres over a wide spectral range. *Electron. Lett.* 15:729.
52. Hsieh, J. J., J. A. Rossi, and J. P. Donnelly. 1976. Room-temperature cw operation of GaInAs/InP double-heterostructure diode lasers emitting at 1.1 μm. *Appl. Phys. Lett.* 28:709.
53. Nelson, R. J., P. D. Wright, P. A. Barnes, R. L. Brown, T. Cella, and R. G. Sobers. 1980. High-output power InGaAsP (1 = 1.3 μm*)* strip-buried hetero-structure lasers. *Appl. Phys. Lett.* 36:358.
54. Hirao, M., S. Tsuji, K. Mizuishi, A. Doi, and M. Nakamura. 1980. Long wavelength InGaAsP/InP lasers for optical fiber communication systems. *J. Opt. Comniun.* 1:l0.
55. Dutta, N. K., and R. J. Nelson. 1982. The case for Auger recombination in $In_{1-x}Ga_xAs_yP_{1-y}$. *J. Appl. Phys.* 53:74.
56. Dutta, N. K., S. G. Napholtz, R. Yen, T. Wessel, T. M. Shen, and N. A. Olsson. 1985. Long wavelength InGaAsP (1 ~ 1.3 μm) modified multiquantum well laser. *Appl. Phys. Lett.* 46:1036.
57. Bates, R. J. S., and S. D. Walker. 1992. Evaluation of all-plastic optical fibre compute data link dispersion limits. *Electran. Left.* 28:996.
58. Gwynne, P. 1996. Digital video disk technology offers increased storage features. *R&D Magazine* 38:40.
59. Yamamoto, S., H. Hayashi, T. Hayakawa, N. Miyauchi, S. Yano, and T. Hijikata. 1982. Room-temperature cw operation in the visible spectral range of 680–700 nm by AlGaAs double heterojunction lasers. *Appl. Phys. Lett.* 41:796.
60. Usui, A., T. Matsumoto, M. Inai, I. Mito, K. Kobayashi, and H. Watanabe. 1985. Room temperature cw operation of visible InGaAsP double heterostructure laser at 671 nm grown by hydride VPE. *Jpn. J. Appl. Phys.* 24:L163.
61. Chong, T. H., and K. Kishino. 1990. Room temperature continuous wave operation of 671-nm wavelength GaInAsP/AlGaAs VSIS lasers. *IEEE Photon. Tech. Lett.* 2:91.
62. Kobayashi, K., S. Kawata, A. Gomyo, I. Hino, and T. Suzuki. 1985. Room-temperature cw operation of AlGaInP double-heterostructure visible lasers. *Electron. Lett.* 21:931.
63. Ikeda, M., Y. Mori, H. Sato, K. Kaneko, and N. Watanabe. 1985. Room-temperature continuous-wave operation of an AlGaInP double heterostructure laser grown by atmospheric pressure metalorganic chemical vapor deposition. *Appl. Phys. Lett.* 47:1027.
64. Ishikawa, M., Y. Ohba, H. Sugawara, M. Yamamoto, and T. Nakanisi. 1986. Room-temperature cw operation of InGaP/InGaAlP visible light laser diodes on GaAs substrates grown by metal-organic chemical vapor deposition. *Appl. Phys. Lett.* 48:207.
65. Hatakoshi, G., K. Nitta, Y. Nishikawa, K. Itaya, and M. Okajima. 1993. High-temperature operation of high-power InGaAlP visible laser. *Proc. SPIE* 1850:388.
66. Hashimoto, J., T. Katsuyama, J. Shinkai, I. Yoshida, and H. Hayashi. 1991. Effects of strained-layer structures on the threshold current density of AlGaInP/GaInP visible lasers. *Appl. Phys. Lett.* 58:879.
67. Honda, S., H. Hamada, M. Shono, R. Hiroyama, K. Yodoshi, and T. Yamaguchi. 1992. Transverse-mode stabilised 630 nm-band AlGaInP strained multiquantum-well laser diodes grown on misoriented ubstrates. *Electron. Lett.* 28:1365.

68. Tanaka, T., H. Yanagisawa, S. Yano, and S. Minagawa. 1993. High temperature operation of 637 nm AlGaInP MQW laser diodes with quaternary QWS grown on misoriented substrates. *Electron. Lett.* 29:24.
69. Ueno, Y., H. Fujii, H. Sawano, K. Kobayashi, K. Hara, A. Gomyo, and K. Endo. 1993. 30-mW 690-nm high-power strained-quantum-well AlGaInP laser. *IEEE J. Quantum Electron.* QE-29:1851.
70. Arimoto, S., M. Yasuda, A. Shima, K. Kadoiwa, T. Kamizato, H. Watanabe, E. Omura, M. Aiga, K. Ikeda, and S. Mitsui. 1993. 150 mW fundamental-transverse-mode operation of 670 nm window laser diode. *IEEE J. Quantum Electmn.* QE-29:1874.
71. Nakayama, N., S. Itoh, K. Nakano, H. Okuyama, M. Ozawa, A. Ishibashi, M. lkeda, and Y. Mori. 1993. Room temperature continuous operation of blue-green laser diodes. *Electron. Lett.* 29:1488.
72. Gaines, J. M., R. R. Drenten, K. W. Haberern, T. Marshall, P. Mensz, and J. Petruzzello. 1993. Blue-green injection lasers contraining pseudomorphic $Zn_{1-x}Mg_xS_ySe_{1-y}$ cladding lasers and operating up to 394 K. *Appl. Phys. Lett.* 62:2462.
73. Haase, M. A., P. F. Baude, M. S. Hagedorn, J. Qiu, J. DePuydt, H. Cheng, S. Guha, G. E. Hofler, and B. J. Wu. 1993. Low-threshold buried-ridge II–VI laser diodes. *Appl. Phys. Lett.* 63:2315.
74. Taniguchi, S., T. Hino, S. Itoh, K. Nakano, N. Nakayama, A. Ishibashi, and M. Ikeda. 1996. 100 h II–VI blue-green laser diode. *Electron. Lett.* 32:552.
75. Nakamura, S., T. Mukai, and M. Senoh. 1991. High-power GaN p–n junction blue-light-emitting diodes. *Jpn. J. Appl. Phys.* 30:L1998.
76. Nakamura, S., M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto. 1996. InGaN MQW structure laser diodes with cleaved mirror facets. *Jpn. J. Appl. Phys.* 35:L217.
77. Itaya, K., M. Onomura, J. Nishio, L. Sugiura, S. Saito, M. Suzuki, J. Rennie, S. Y. Nunoue, M. Yamamoto, H. Fujimoto, Y. Kokubun, Y. Ohba, G. Hatakoshi, and M. Ishikawa. 1996. Room temperature pulsed operation of nitride based multi-quantum-well laser diodes with cleaved facets on conventional c-face sapphire substrates. *Jpn. J. Appl. Phys.* (Part 2) 35:L1315.
78. Akasaki, I., S. Sota, H. Sakai, T. Tanaka, M. Koike, and H. Amano. 1996. Shortest wavelength semiconductor laser diode. *Electron. Lett.* 32:1105.
79. Nakamura, S., M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto. 1996. InGaN multi-quantum-well structure laser diodes grown on $MgAl_2O_4$ substrates. *Appl. Phys. Lett.* 68:2105.
80. Akasaki, I., and H. Amano. 1996, November. Progress and future prospects of group III nitride semiconductors. *LEOS'96,* Plen2, Boston.
81. Nakamura, S., M. Senoh, S. I. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, Y. Sugimoto, and H. Kiyoku. 1996, November. First room-temperature continuous-wave operation of InGaN multi-quantum-well-structure laser diodes. *LEOS'96,* PD1.l, Boston.
82. Hasnain, G., K. Tai, J. D. Wynn, Y. H. Wang, R. J. Fischer, M. Hong, B. E. Weir, G. J. Zydzik, J. P. Mannaerts, J. Gamelin, and A. Y. Cho. 1990. Continuous wave top surface emitting quantum well lasers using hybrid metal/semiconductor reflectors. *Electron. Lett.* 26:1590.
83. Morgan, R. A., M. K. Hibbs-Brenner, T. M. Marta, R. A. Walterson, S. Bounnak, E. L. Kalweit, and J. A. Lehman. 1995. 200 degrees-C. 96-nm wavelength range, continuous-wave lasing from unbonded GaAs MOVPE-grown vertical cavity surface-emitting lasers. *IEEE Photon. Tech. Lett.* 7:441.
84. Zhou, P., J. L. Cheng, C. E Schaus, S. Z. Sun, K. Zheng, E. Armour, C. Hains, W. Hsin, D. R. Myers, and G. A. Vawter. 1991. Low series resistance high-efficiency GaAs/AlGaAs vertical-cavity surface-emitting lasers with continuously graded mirrors grown by MOCVD. *IEEE Photon. Tech. Lett.* 3:591.
85. Tan, M. R. T., K. H. Hahn, Y. M. D. Houng, and S. Y. Wang, 1995. Surface emitting laser for multimode data link applications. *HP J.,* February:67.

86. Tai, K., L. Yang, Y. H. Wang, J. D. Wynn, and A. Y. Cho. 1990. Drastic reduction of series resistance in doped semiconductor distributed Bragg reflectors for surface-emitting lasers. *Appl. Phys. Lett.* 56:2496.
87. Zhou, P., J. Cheng, C. E Schaus, S. Z. Sun, K. Zheng, E. Armour, C. Hains, W. Hsin, D. R. Myers, and G. A. Vawter. 1991. Low series resistance high-efficiency GaAs AlGaAs vertical-cavity surface-emitting lasers with continuously graded mirrors grown by MOCVD. *IEEE Photon. Tech. Lett.* 3591.
88. Schubert, E. F. L., W. Tu, G. J. Zydzik, R. F. Kopf, A. Benvenuti, and M. R. Pinto. 1992. Elimination of heterojunction band discontinuities by modulation doping. *Appl. Phys. Lett.* 60:466.
89. Peters, M. G., D. B. Young, F. H. Peters, J. W. Scott, B. J. Thibeault, and L. A. Coldren. 1994. 17.3-percent peak wall plug efficiency vertical-cavity surface-emitting lasers using lower barrier mirrors. *IEEE Photon. Tech. Lett.* 6:31.
90. Peters, M. G., B. J. Thibeault, D. B. Young, J. W. Scott, F. H. Peters, A. C. Gossard, and L. A. Coldren. 1993. Band-gap engineered digital alloy interfaces for lower resistance vertical-cavity surface-emitting lasers. *Appl. Phys. Lett.* 63:3411.
91. Lear, K. L., S. A. Chalmers, and K. P. Killeen. 1993. Low threshold voltage vertical cavity surface-emitting laser. *Electron. Lett.* 29:584.
92. Young, D. B., J. W. Scott, F. H. Peters, M. G. Peters, M. L. Majewski, B. J. Thibeault, S. W. Corzine, and L. A. Coldren. 1993. Enhanced performance of offset-gain high-barrier vertical-cavity surface-emitting lasers. *IEEE J. Quantum Electron.* QE-29:2013.
93. Schubert, E. F., A. Fischer, Y. Horikoshi, and K. Ploog. 1985. GaAs sawtooth superlattice laser emitting at wavelength $1 > 0.9\,\mu m$. *Appl. Phys. Lett.* 47:219.
94. Kojima, K., R. A. Morgan, T. Mullaly, G. D. Guth, M. W. Focht, R. E. Leibenguth, and M. T. Asom. 1993. Reduction of *p*-doped mirror electrical resistance of GaAdAlGaAs vertical-cavity surface-emitting lasers by delta doping. *Electron. Lett.* 29:1771.
95. Scott, J. W., B. J. Thibeault, D. B. Young, L. A. Coldren, and F. H. Peters. 1994. High efficiency submilliamp vertical cavity lasers with intracavity contacts. *IEEE Photon. Tech. Lett.* 6:678.
96. Rochus, S., M. Hauser, T. Rohr, H. Kratzer, G. Böhm, W. Klein, G. Triinkle, and G. Weimann. 1995. Submilliamp vertical-cavity surface-emitting lasers with buried lateral-current confinement. *IEEE Photon. Tech. Lett.* 7:968.
97. Morgan, R. A., M. K. Hibbs-Brenner, J. A. Lehman, E. L. Kaiweit, R. A. Walterson, T. M. Marta, and T. Akinwande. 1995. Hybrid dielectric/AlGaAs mirror spatially filtered vertical cavity top-surface emitting laser. *Appl. Phys. Lett.* 66:1157.
98. Dudley, J. J., D. L. Crawford, and J. E. Bowers. 1992. Temperature dependence of the properties of DBR mirrors used in surface normal optoelectronic devices. *IEEE Photon. Tech. Lett.* 4:311.
99. Lebby, M., C. A. Gaw, W. B. Jiang, P. A. Kiely, P. R. Claisse, and J. Ramdani. 1996. Vertical-cavity surface-emitting lasers for communication applications. OSA annual meeting, WR1, October, Rochester, N.Y.
100. Lee, Y. H., B. Tell, K. F. Brown-Goebeler, R. E. Leibenguth, and V. D. Mattera. 1991. Deep-red CW top surface-emitting vertical-cavity AlGaAs superlattice lasers. *IEEE Photon. Tech. Lett.* 3:108.
101. Shin, H. E., Y. G. Ju, J. H. Shin, J. H. Ser, T. Kim, E. K. Lee, I. Kim, and Y. H. Lee. 1996. 780 nm oxidised vertical-cavity surface-emitting lasers with $Al_{0.1}Ga_{0.89}As$ quantum wells. *Electron. Lett.* 32:1287.
102. Kim, T., T. K. Kim, E. K. Lee, J. Y. Kim, and T. I. Kim. 1995. A single transverse mode operation of top surface emitting laser diode with an integrated photo-diode. *Pmc. LEOS'95* 2:416.
103. Schneider, R. P., Jr., K. D. Choquette, J. A. Lott, K. L. Lear, J. J. Figiel, and K. J. Malloy. 1994. Efficient room-temperature continuous-wave AlGaInP/AlGaAs visible (670 nm) vertical-cavity surface-emitting laser diodes. *IEEE Photon. Tech. Lett.* 6:313.

104. Choquette, K. D., R. P. Schneider, M. H. Crawford, K. M. Geib, and J. J. Figiel. 1995. Continuous wave operation of 640–660 nm selectively oxidised AlGaInP vertical-cavity lasers. *Electron. Len.* 31:1145.
105. Schneider, R. P., Jr., M. H. Crawford, K. D. Choquette, K. L. Lear, S. P. Kilcoyne, and J. J. Figiel. 1995. Improved AlGaInP-based red (670–690 nm) surface-emitting lasers with novel C-doped short-cavity epitaxial design. *Appl. Phys. Lett.* 67:329.
106. Crawford, M. H., R. P. Schneider, Jr., K. D. Choquette, and K. L. Lear. 1995. Temperature-dependent characteristics and single-mode performance of AlGaInP-based 670–690-nm vertical-cavity surface-emitting lasers. *IEEE Photon. Tech. LRtt.* 7:724.
107. Baba, T., Y. Yogo, K. Suzuki, F. Koyama, and K. Iga. 1993. Near room temperature continuous wave lasing characteristics of GaInAsP/InP surface emitting laser. *Electron. Lett.* 29:913.
108. Dudley, J. J., M. Ishikawa, B. I. Miller, D. I. Babic, R. Mirin, W. B. Jiang, J. E. Bowers, and E. L. Hu. 1992. 144°C operation of 1.3 μm InGaAsP vertical cavity lasers on GaAs substrates. *Appl. Phys. Lett.* 61:3095.
109. Dudley, J. J., D. I. Babic, R. Mirin, L. Yang, B. I. Miller, R. J. Ram, T. Reynolds, E. L. Hu, and J. E. Bowers. 1994. Low threshold, wafer fused long wavelength vertical cavity lasers. *Appl. Phys. Lett.* 64:1463.
110. Shin, H. K., I. Kim, E. J. Kim, J. H. Kim, E. K. Lee, M. K. Lee, J. K. Mun, C. S. Park, and Y. S. Yi. 1996. Vertical-cavity surface-emitting lasers for optical data storage. *Jpn. J. Appl. Phys.* (Part l), 35:506.
111. Hasnain, G., and K. Tai. 1992. Self-monitoring semiconductor laser device. U.S. Patent No. 5,136,603.
112. Hasnain, G., K. Tai, Y. H. Wang, J. D. Wynn, K. D. Choquette, B. E. Weir, N. K. Dutta, and A. Y. Cho. 1991. Monolithic integration of photodetector with vertical cavity surface emitting laser. *EZectron. Lett.* 27:1630.
113. Hibbs-Brenner, M. K. 1995. Integrated laser power monitor. U.S. Patent No. 5,475,701.
114. Bjork, G., and Y. Yamamoto. 1991. Analysis of semiconductor microcavity lasers using rate equations. *IEEE J. Quantum Electron.* QE-27:2386.
115. Ram, R. J., E. Goobar, M. G. Peters, L. A. Coldren, and J. E. Bowers. 1996. Spontaneous emission factor in post microcavity lasers. *IEEE Photon. Tech. Lett.* 8:599.
116. Huffaker, D. L., D. G. Deppe, and K. Kumar. 1994. Native-oxide ring contact for low threshold vertical-cavity lasers. Appl. *Phys. Lett.* 65:97.
117. MacDougal, M. H., P. Daniel Dapkus, V. Pudikov, H. M. Zhao, and G. M. Yang. 1995. Ultralow threshold current vertical-cavity surface-emitting lasers with AlAs oxide-GaAs distributed Bragg reflectors. *IEEE Photon. Tech. Lett.* 7:229.
118. MacDougal, M. H., G. M. Yang, A. E. Bond, C. K. Lin, D. Tishinin, and P. D. Dapkus. 1996. Electrically-pumped vertical-cavity lasers with Al_xO_yGaAs reflectors. *IEEE Photon. Tech. Lett.* 8:310.
119. Lear, K. L., A. Mar, K. D. Choquette, S. P. Kilcoyne, R. P. Schneider, Jr., and K. M. Geib. 1996. High frequency modulation of oxide-confined vertical cavity surface emitting lasers. *Electron. Lett.* 32:457.
120. Thibeault, B. J., K. Bertilsson, E. R. Hegblom, E. Strzelecka, P. D. Floyd, R. Naone, and L. A. Coldren. 1997. High-speed characteristics of low-optical loss oxide-apertured vertical-cavity lasers. *IEEE Photon. Tech. Lett.* 9:11.

This page intentionally left blank

6

Detectors for Fiber Optics

Carolyn J. Sher DeCusatis
Department of Electrical and Computer Engineering, State University of New York at New Paltz, New Paltz, NY 12561

Ching-Long (John) Jiang
Amp Incorporated, Lytel Division, Somerville, New Jersey 08876

6.1. DETECTOR TERMINOLOGY AND CHARACTERISTICS

Every detector specification should include a picture and/or physical description of the part, including dimensions and construction (i.e., plastic housing). In this section we have tried to be inclusive in our list of terms, which means that not all of these quantities will apply to every detector specification. Since specifications are not standardized, it is impossible to include all possible terms used; however, most detectors are described by certain standard figures of merit, which will be discussed in this section. It is important to consider the manufacturer's context for all values; a detector designed for a specific application may not be appropriate for a different application, even though the specification seems appropriate.

Among the figures of merit used to characterize the performance of different detectors is responsivity, or response—the sensitivity of the detector to input flux. It is given by

$$R(\lambda) = I/\phi\ (\lambda) \tag{6.1}$$

where I is the detector output signal (in amps) and ϕ is the incident light signal on the detector (in watts). Thus, the units of responsivity are amps per watt. Even when the detector is not illuminated, some current will flow; this dark current may be subtracted from the detector output signal when determining detector performance. Dark current is the thermally generated current in a photodiode under a completely dark environment; it depends on the material, doping, and structure of the photodiode. It is the lowest level of thermal noise. Dark current

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

in photodiodes limits the sensitivity (minimum detectable power). The reduction of dark current is important for the improvement of minimum detectable power. It is usually simply measured and then subtracted from the flux, like background, in most specifications. However, the dark current is temperature dependent, so care must be taken to evaluate it over the expected operating conditions. It is not a good idea for the anticipated signal to be a small fraction of the dark current; root mean square (rms) noise in the dark current may mask the signal. Responsivity is defined at a specific wavelength; the term *spectral responsivity* is used to describe the variation at different wavelengths. Responsivity versus wavelength is often included in a specification as a graph, as well as placed in a performance chart at a specified wavelength.

Quantum efficiency (QE) is the ratio of the number of electron-hole pairs collected at the terminals to the number of photons in the incident light. It depends on the material from which the detector is made and is determined primarily by reflectivity, absorption coefficient, and carrier diffusion length. As the absorption coefficient is dependent on the incident light wavelength, the quantum efficiency has a spectral response. Quantum efficiency is the fundamental efficiency of the diode for converting photons into electron-hole pairs.

For example, the quantum efficiency of a PIN diode can be calculated by

$$QE = (1 - R)T(1 - e^{-\alpha W}) \tag{6.2}$$

where R is the surface reflectivity, T is the transmission of any lossy electrode layers, W is the thickness of the absorbing layer, and α is the absorption coefficient.

Quantum efficiency affects detector performance through the responsivity (R), which can be calculated from quantum efficiency:

$$R(\lambda) = QE\ \lambda\ q/h\ c \tag{6.3}$$

where q is the charge of an electron (1.6×10^{-19} coulomb), λ is the wavelength of the incident photon, h is Planck's constant (6.626×10^{-34} W), and c is the velocity of light (3×10^{8} m/s). If wavelength is in nanometers and R is responsivity flux, then the units of responsivity are amperes per watt. Responsivity is the ratio of the diode's output current to input optical power and is given in amperes per watt (A/W). A PIN photodiode typically has a responsivity of 0.6 to 0.8 A/W. A responsivity of 0.8 A/W means that incident light having 50 microwatts of power results in 40 microamps of current; in other words,

$$I = 50\,\mu W \times 0.8\,A/W = 40\,\mu A \tag{6.4}$$

where I is the photodiode current. For an avalanche photodiode (APD), a typical responsivity is 80 A/W. The same 50 microwatts of optical power now produces 4 mA of current:

$$I = 50\,\mu W \times 80\,A/W = 4\,mA \tag{6.5}$$

The minimum power detectable by the photodiode determines the lowest level of incident optical power that the photodiode can detect. It is related to the dark current in the diode, since the dark current will set the lower limit. Other noise sources are factors, including those associated with the diode and those associated with the receiver. The noise floor of a photodiode, which tells us the minimum detectable power, is the ratio of noise current to responsivity:

$$\text{Noise floor} = \text{noise/responsivity} \tag{6.6}$$

For initial evaluation of a photodiode, we can use the dark current to estimate the noise floor. Consider a photodiode with R = 0.8 A/W and a dark current of 2 nA. The minimum detectable power is

$$\text{Noise floor} = (2\,nA)/(0.8\,nA/nW) = 2.5\,nW \tag{6.7}$$

More precise estimates must include other noise sources, such as thermal and shot noise. As discussed, the noise depends on current, load resistance, temperature, and bandwidth.

Response time is the time required for the photodiode to respond to an incoming optical signal and produce an external current. Similarly to a source, response time is usually specified as a rise time and a fall time, measured between the 10% and 90% points of amplitude (other specifications may measure rise and fall times at the 20%–80% points, or when the signal rises or falls to 1/e of its initial value). The bandwidth of a photodiode can be limited by either its rise time and fall time or its RC time constant, whichever results in the slower speed or bandwidth. The bandwidth of a circuit limited by the RC time constant is

$$B = 1/2\pi RC \tag{6.8}$$

where R is the load resistance and C is the diode capacitance. Fig 6.1 shows the equivalent circuit model of a photodiode. It consists of a current source in parallel with a resistance and a capacitance. It appears as a low-pass filter, a resistor-capacitor network that passes low frequencies and attenuates high frequencies. The cutoff frequency, which is the frequency that is attenuated 3 dB, marks the

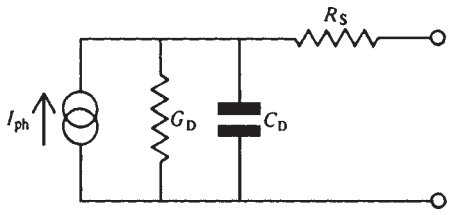


Figure 6.1 Small-signal equivalent circuit for a reversed biased photodiode.

3-dB bandwidth. Photodiodes for high-speed operation must have a very low capacitance. The capacitance in a photodiode is mainly the junction capacitance formed at the pn junction, as well as any capacitance contributed by the packaging.

Bias voltage refers to an external voltage applied to the detector and will be more fully described in the following section. Photodiodes require bias voltages ranging from as low as 0 V for some PIN photodiodes to several hundred volts for APDs. Bias voltage significantly affects operation, since dark current, responsivity, and response time all increase with bias voltage. APDs are usually biased near their avalanche breakdown point to ensure fast response.

Active area and effective sensing area are just what they sound like: the size of the detecting surface of the detection element. (This figure of merit is important to consider when modifying a single-mode detector for use on multimode fiber.) The uniformity of response refers to the percentage change of the sensitivity across the active area. Operating temperature is the temperature range over which a detector is accurate and will not be damaged by being powered. However, changes in sensitivity and dark current must be taken into account: read the manual. Storage temperature will have a considerably larger range; basically, it describes the temperature range under which the detector will not melt, freeze, or otherwise be damaged or lose its operating characteristics.

NEP, or noise equivalent power, is the amount of flux that would create a signal of the same strength as the rms detector noise. In other words, it is a measure of the minimum detectable signal. For this reason, it is the most commonly used version of the more generic figure of merit, noise equivalent detector input. More formally, it may be defined as the optical power (of a given wavelength or spectral content) required to produce a detector current equal to the root mean square (rms) noise in a unit bandwidth of 1 Hz:

$$\mathrm{NEP}\ (\lambda) = i_n\ (\lambda)/R(\lambda) \tag{6.9}$$

where i_n is the rms noise current and R is the responsivity, defined previously. It can be shown [2] that to a good approximation,

$$\mathrm{NEP} = 2\,h\,c/QE\ \lambda \tag{6.10}$$

where this expression gives the NEP of an ideal diode when QE = 1. If the dark current is large, this expression may be approximated by

$$\mathrm{NEP} = h\ c\ (2\ q\ I)^{1/2}/QE\ q\ \lambda \tag{6.11}$$

where I is the detector current. Sometimes it is easier to work with detectivity, which is the reciprocal of NEP. The higher the detectivity, the smaller the signal a detector can measure; this is a convenient way to characterize more sensitive detectors. Detectivity and NEP vary with the inverse of the square of active area of the detector, as well as with temperature, wavelength, modulation frequency, signal voltage, and bandwidth. For a photodiode detecting monochromatic light and dominated by dark current, detectivity is given by

$$D = QE\ q\ \lambda/h\ c\ (2\ q\ I)^{1/2} \tag{6.12}$$

The quantity-specific detectivity accounts for the fact that dark current is often proportional to detector area, A; it is defined by

$$D^* = D\ A^{1/2} \tag{6.13}$$

Normalized detectivity is detectivity multiplied by the square root of the product of active area and bandwidth; this product is usually constant and allows comparison of different detector types independent of size and bandwidth limits. This is because most detector noise is white noise (Gaussian power spectra), and white noise power is proportional to the bandwidth of the detector electronics. Thus the noise signal is proportional to the square root of bandwidth. Also, note that electrical noise power is usually proportional to detector area and the voltage that provides a measure of that noise is proportional to the square root of power. Normalized detectivity is given by:

$$D_n = D\ (A\ B)^{1/2} = (A\ B)^{1/2}/NEP \tag{6.14}$$

where B is the bandwidth. The units are $(cm\ Hz)^{1/2}\ W^{-1}$. Normalized detectivity is a function of wavelength and spectral responsivity; it is often quoted as normalized spectral responsivity.

Bandwidth, B, is the range of frequencies over which a particular instrument is designed to function within specified limits. Bandwidth is often adjusted to limit noise; in some specifications it is chosen as 1 Hz, so NEP is quoted in watts/Hz. Wide-bandwidth detectors required in optical datacom often operate into a low resistance and require a minimal signal current much larger than the dark current; the load resistance, amplifier, and other noise sources can make the use of NEP, D, D*, and D_n inappropriate for characterizing these applications.

Linearity range is the range of incident radiant flux over which the signal output is a linear function of the input. The lower limit of linearity is NEP, and the upper limit is saturation. Saturation occurs when the detector begins to form less signal output for the same increase of input flux. When a detector begins to saturate, it has reached the end of its linear range. Dynamic range can be used to describe nonlinear detectors, like the human eye. Although datacom systems do not typically use filters on the detector elements, neutral density filters can be used to increase the dynamic range of a detector system by creating islands of linearity, whose actual flux is determined by dividing output signals of the detector by the transmission of the filter.

Without filtering, the dynamic range would be limited to the linear range of the detector, which would be less because the detector would saturate without the filter to limit the incident flux. The units of linear range are incident radiant flux or power (watts or irradiance). Measuring the response of a detector to flux is known as calibration. Some detectors can be self-calibrated, whereas others require manufacturer calibration. Calibration certificates are supplied by most

manufacturers for fiber-optic test instrumentation; they are dated and have certain time limits. The gain, also known as the amplification, is the ratio of electron-hole pairs generated per incident photon. Sometimes detector electronics allows the user to adjust the gain. Wiring and pin output diagrams tell the user how to operate the equipment, by schematically showing how to connect the input and output leads.

6.2. PIN PHOTODIODE

Photodiode detectors used in data communications are solid-state devices; to understand their function, we must first describe a bit of semiconductor physics. For the interested reader, other introductory references to solid-state physics, semiconductors, and condensed matter are available [2]. In a solid-state device, the electron potential can be described in terms of conduction bands and valence bands, rather than individual potential wells. The highest energy level containing electrons is called the Fermi level. If a material is a conductor, the conduction and valence bands overlap and charge carriers (electrons or holes) flow freely; the material carries an electrical current. An insulator is a material for which there is a large enough gap between the conduction and valence bands to prohibit the flow of carriers; the Fermi level lies in the middle of the forbidden region between bands, called the bandgap. A semiconductor is a material for which the bandgap is small enough that carriers can be excited into the conduction band with some stimulus; the Fermi level lies at the edge of the valence band (if the majority of carriers are holes) or the edge of the conduction band (if the majority of carriers are electrons). The first case is called a p-type semiconductor, the second is called n-type. These materials are useful for optical detection because incident light can excite electrons across the bandgap and generate a photocurrent.

The simplest photodiode is the pn photodiode. Although this type of detector is not widely used in fiber optics, it serves the purpose of illustrating the basic ideas of semiconductor photodetection, since other devices—the Positive-intrinsic-negative (PIN) and avalanche photodiodes—are designed to overcome the limitations of the pn diode. When the pn photodiode is reverse biased (negative battery terminal connected to p-type material), very little current flows. The applied electric field creates a depletion region on either side of the pn junction. Carriers—free electrons and holes—leave the junction area. In other words, electrons migrate toward the negative terminal of the device and holes toward the positive terminal. Because the depletion region has no carriers, its resistance is very high, and most of the voltage drop occurs across the junction. As a result, electrical fields are high in this region and negligible elsewhere. An incident photon absorbed by the diode gives a bound electron sufficient energy to move from the valence band to the conduction band, creating a free electron and a hole. If this creation of carriers occurs in the depletion region, the carriers quickly

separate and drift rapidly toward their respective regions. This movement sets an electron flowing as current in the external circuit.

The structure of the PIN diode is designed to overcome the deficiencies of its pn counterpart. The PIN diode is a photoconductive device formed from a sandwich of three layers of crystal, each layer with different band structures caused by adding impurities (doping) to the base material, usually indium gallium arsenide, silicon, or germanium. The layers are doped in this arrangement: p-type (or positive) on top, intrinsic, meaning undoped, in a thin middle layer, and n-type (or negative) type on the bottom. For a silicon crystal a typical p-type impurity would be boron, and indium would be a p-type impurity for germanium [2–6]. Actually, the intrinsic layer may also be lightly doped, though not enough to make it either p-type or n-type. The change in potential at the interface has the effect of influencing the direction of current flow, creating a diode. Obviously, the name PIN diode comes from the sandwich of p-type, intrinsic, and n-type layers.

The structure of a typical PIN photodiode is shown in Fig. 6.2. The p-type and n-type silicon form a potential at the intrinsic region; this potential gradient depletes the junction region of charge carriers, both electrons and holes, and results in the conduction band bending. The intrinsic region has no free carriers, and thus exhibits high resistance. The junction drives holes into the p-type material and electrons into the n-type material. The difference in potential of the two

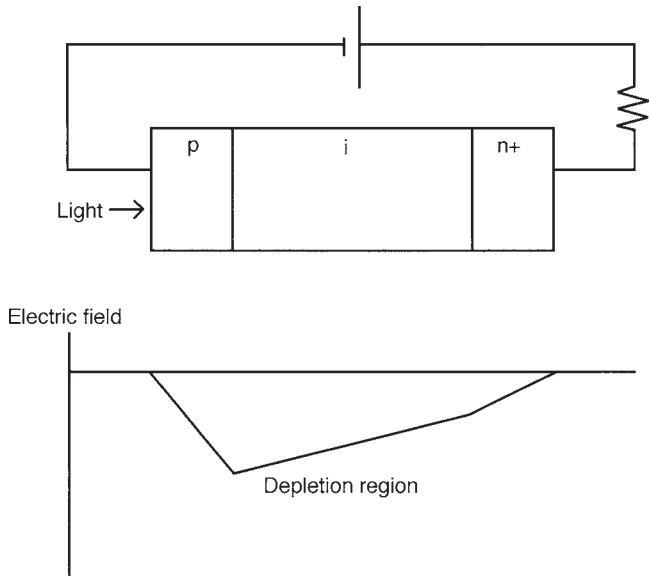


Figure 6.2 PIN diode.

materials determines the energy an electron must have to flow through the junction. When photons fall on the active area of the device, they generate carriers near the junction, resulting in a voltage difference between the p-type and n-type regions. If the diode is connected to external circuitry, a current will flow that is proportional to the illumination. The PIN diode structure addresses the main problem with pn diodes, namely, providing a large depletion region for the absorption of photons. There is a tradeoff involved in the design of PIN diodes. Since most of the photons are absorbed in the intrinsic region, a thick intrinsic layer is desirable to improve photon-carrier conversion efficiency (to increase the probability of a photon being absorbed in the intrinsic region). On the other hand, a thin intrinsic region is desirable for high-speed devices, since it reduces the transit time of photogenerated carriers. These two conditions must be balanced in the design of PIN diodes.

Photodiodes can be operated either with or without a bias voltage. Unbiased operation is called the photovoltaic mode; certain types of noise, including 1/f noise, are lower and the NEP is better at low frequencies. Signal-to-noise ratio is superior to the biased mode of operation for frequencies below about 100 kHz [6]. Biasing (connecting a voltage potential to the two sides of the junction) will sweep carriers out of the junction region faster and change the energy requirement for carrier generation to a limited extent. Biased operation (photoconductive mode) can be either forward or reverse biased. The reverse bias of the junction (positive potential connected to the n-side and negative connected to the p-side) reduces junction capacitance and improves response time; for this reason it is the preferred operation mode for pulsed detectors. A PIN diode used for photodetection may also be forward biased (the positive potential connected to the p-side and the negative to the n-side of the junction), to make the potential scaled for current to flow less, or in other words to increase the sensitivity of the detector (Fig. 6.3).

An advantage of the PIN structure is that the operating wavelength and voltage, diode capacitance, and frequency response may all be predetermined during the manufacturing process. For a diode whose intrinsic layer thickness is w with an applied bias voltage of *V*, the self-capacitance of the diode, C, approaches that of a parallel plate capacitor,

$$C = \varepsilon_o \, \varepsilon_1 \, A_o / w \tag{6.15}$$

where A_o is the junction area, ε_o the free space permittivity (8.849×10^{-12} farads/m), and ε_1 the relative permittivity. Taking typical values of $\varepsilon_1 = 12$, w = 50 microns and $A_o = 10^{-7}\,m^2$, C = 0.2 pF. Quantum efficiencies of 0.8 or higher can be achieved at wavelengths of 0.8–0.9 micron, with dark currents less than 1 nA at room temperature. Some typical responsivities for common materials are given in Table 6.1.

The sensitivity of a PIN diode can vary widely by quality of manufacture. A typical PIN diode size ranges from 5 mm × 5 mm to 25 mm × 25 mm. Ideally, the

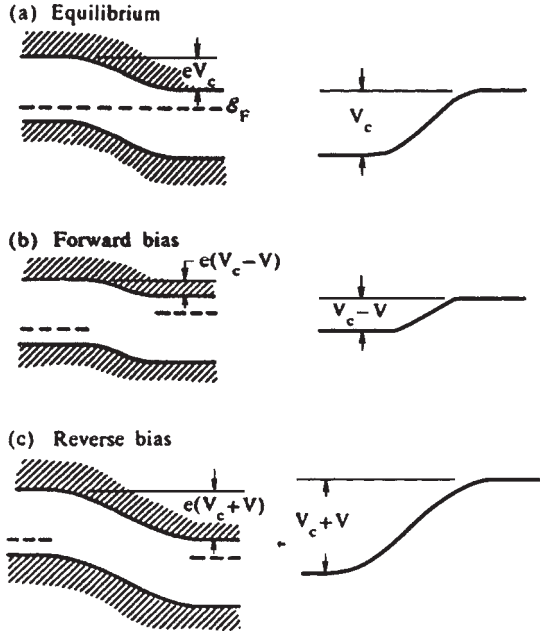


Figure 6.3 Forward and reverse bias of a diode [2].

Table 6.1
Common Semiconductor PIN Diode Properties.

Material	Wavelength range (μm)	Peak responsivity (A/W)
Silicon	0.3–1.1	0.5 A/W at 0.8 μm
Germanium	0.5–1.8	0.8 A/W at 0.7 μm
InGaAs	1.0–1.7	1.1 A/W at 1.7 μm

detection surface will be uniformly sensitive (a detector profiler at the National Institute for Standards and Technology, by using extremely well-focused light sources, can determine the sensitivity of a detector's surface [3]). Most applications require that the detector be uniformly illuminated or overfilled. The spectral responsivity of an uncorrected silicon photodiode is shown in Fig. 6.4. The typical QE curve is also shown for comparison. An ideal silicon detector would have zero responsivity and QE for photons whose energies are less than the bandgap, or wavelengths much longer than about 1.1 micron. Just below the long wave limit, this ideal diode would have 100% QE and responsivity close to 1 A/W;

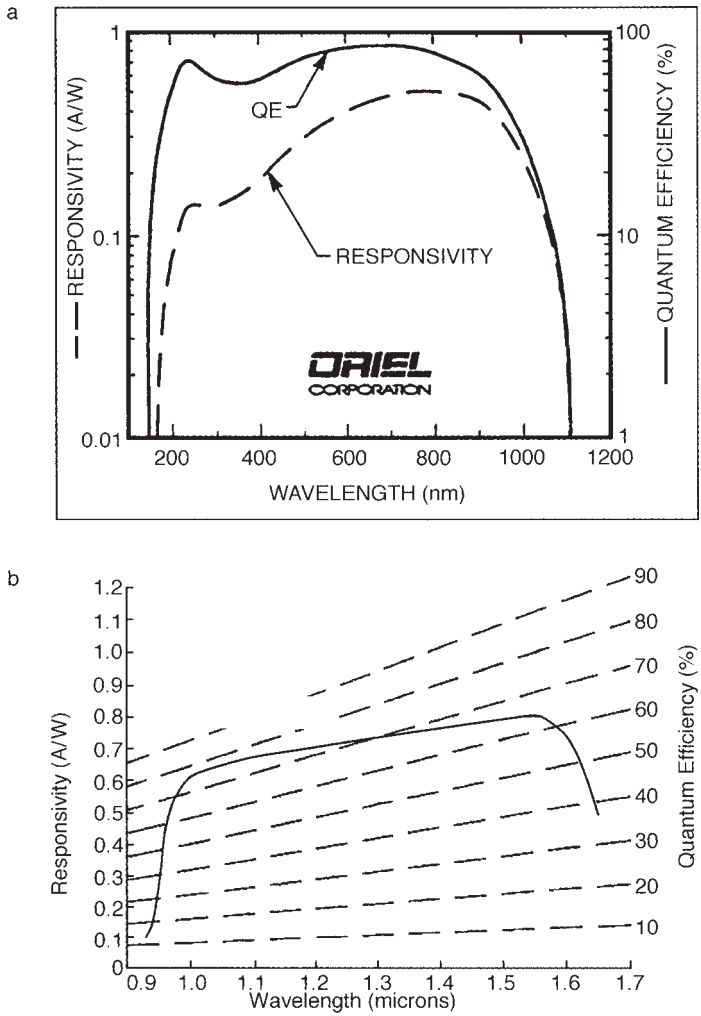


Figure 6.4 PIN diode spectral response and quantum efficiency. (a) silicon; (b) InGaAs.

responsivity vs. wavelength would be expected to follow the intrinsic spectral response of the material. In practice, this does not happen; these detectors are less sensitive in the blue region, which can sometimes be enhanced by clever doping, but not more than an order of magnitude. This lack of sensitivity exists because there are fewer short wavelength photons per watt, so that responsivity in terms of power drops off, and because more energetic blue photons may not be absorbed

in the junction region. For color-sensitive applications such as photometry, filters are used so that detectors will respond photometrically, or to the standardized CIE color coordinates; however, the lack of overall sensitivity in the blue region can potentially create noise problems when measuring a low-intensity blue signal. In the deep ultraviolet, photons are often absorbed before they reach the sensitive region by detector windows or surface coatings on the semiconductor. The departure from 100% QE in real devices is typically due to Fresnel reflections from the detector surface. The long-wavelength cutoff is more gradual than expected for an ideal device because the absorption coefficient decreases at long wavelengths, so more photons pass though the photosensitive layers and do not contribute to the QE. As a result, QE tends to roll off gradually near the bandgap limit. This response is typical for silicon devices, which make excellent detectors in the wavelength range 0.8–0.9 micron. A common material for fiber-optic applications is InGaAs, which is most sensitive in the near infrared (0.8 to 1.7 microns). Other PIN diode materials include HgCdZnTe for wavelengths of 2 to 12 microns. In the 1940s, a popular photoconductive material for infrared solid-state detectors was introduced, lead sulfide (PbS); this material is still commonly used in the region from 1–4 microns [6].

The speed of the detector depends on the junction capacitance. Two of the most widely used methods of reducing junction area use are mesa structures and planar structures with selective diffusion of the contact area. Another method is increasing the thickness of the depletion region, but this increases the transit time of photogenerated carriers. Strategies to improve transit time include edge-illuminated structure and absorptive resonance in the i-layer (see Section 6.3.5) [7].

PIN diode detectors are not very sensitive to temperature (−25°C to +80°C) or shock and vibration, making them an ideal choice for a data communications transceiver. It is very important to keep the surface of any detector clean. This becomes an issue with PIN diode detectors because they are sufficiently rugged that they can be brought into applications that expose them to contamination. Both transceivers and optical connectors should be cleaned regularly during use to avoid dust and dirt buildup.

A sample specification for a photodiode is given in Table 6.2. Bias voltage (V) is the voltage applied to a silicon photodiode to change the potential photoelectrons must scale to become part of the signal. Bias voltage is basically a set operating characteristics in a prepackaged detector, in this case −24 V. Shunt resistance is the resistance of a silicon photodiode when not biased. Junction capacitance is the capacitance of a silicon photodiode when not biased. Breakdown voltage is the voltage applied as a bias that is large enough to create signal on its own. When this happens, the contribution of photoelectrons is minimal, so the detector cannot function. However, once the incorrect bias is removed, the detector should return to normal.

Table 6.2
Sample Specification for a PIN Diode.

Receiver Section						
Parameter	Symbol	Test Conditions	Min.	Typical	Max.	Units
Data rate (NRZ)	B	—	10	—	156	Mb/s
Sensitivity (avg)	P_{OH}	62.5 µm fiber .275 NA, BER $\leq 10^{-10}$	–32.5	—	–14.0	dBm
Optical wavelength	λ_{IN}	—	1270	—	1380	nm
Duty cycle	—	—	25	50	75	%
Output risetime	t_{TLH}	20–80% 50Ω to V_{CC}-2V	.5	—	2.5	ns
Output falltime	t_{THL}	80–20% 50Ω to V_{CC}-2V	.5	—	2.5	ns
Output voltage	V_{OL}	—	V_{CC}-1.025	—	V_{CC}-.88	V
	V_{OH}		V_{CC}-1.81	—	V_{CC}-1.62	V
Signal detect	V_A	$P_{IN} > P_A$	V_{CC}-1.025	—	V_{CC}-88	V
	V_D	$P_{IN} < P_D$	V_{CC}-1.81	—	V_{CC}-1.62	V
P_{IN} power levels:						
Deassert	P_B	—	–39.0 or P_B	—	–32.5	dBm
Assert	P_A		–38.0	—	–30.0	dBm
Hysteresis	—	—	1.5	2.0	—	dB
Signal detect delay time:						
Deassert	—	—	—	—	50	µS
Assert	—	—	—	—	50	µS
Power supply voltage	V_{CC}-V_{EE}	—	4.75	5.0	5.25	V
Power supply current	I_{CC} or I_{EE}	—	—	—	150	mA
Operating temperature	T_A	—	0	—	70	°C

Absolute Maximum Ratings: Transceiver						
Parameter	Symbol	Test Conditions	Min.	Typical	Max.	Units
Storage temperature	—	—	–40	—	100	°C
Lead soldering limits	—	—	—	—	240/10	°C/s
Supply voltage	V_{CC}-V_{EE}	—	–.2	—	7.00	V

To increase response speed and quantum efficiency, a variation on the PIN diode, known as the Schottky barrier diode, can be used. This approach will treated in a later section (6.3.3). For wavelengths longer than about 0.8 micron, a heterojunction diode may be used for this same reason. Heterojunction diodes retain the PIN sandwich structure, but the surface layer is doped to have a wider

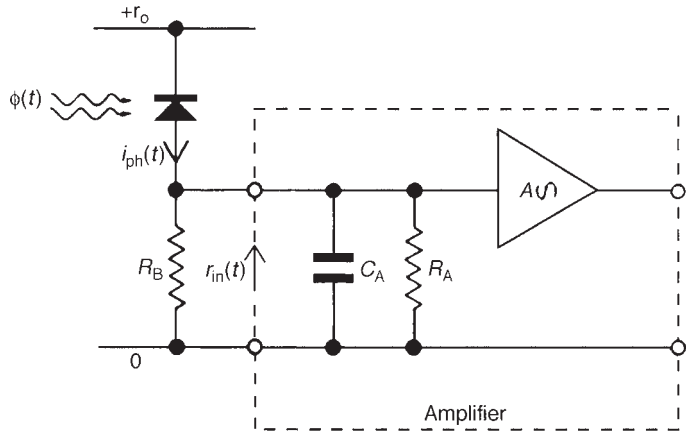


Figure 6.5 PIN diode and amplifier equivalent circuit.

bandgap and thus reduced absorption. In this case, absorption is strongest in the narrower bandgap region at the heterojunction, where the electric field is maximum; hence, good quantum efficiencies can be obtained. The most common material systems for heterojunction diodes are InGaAsP on an InP substrate, or GaAlAsSb on a GaSb substrate.

A typical circuit for biased operation of a photodiode is shown in Fig. 6.5, where C_a and R_a represent the input impedance of a postamplifier. The photodiode impulse response will differ from an ideal square wave for several reasons, including transit time resulting from the drift of carriers across the depletion layer, delay caused by the diffusion of carriers generated outside the depletion layer, and the RC time constant of the diode and its load. If we return to the photodiode equivalent circuit of Fig. 6.1 and insert it into this typical bias circuit, we can make the approximation that R_s is much smaller than R_a to arrive at the equivalent small signal circuit model of Fig. 6.6. In this model the resistance R is approximately equal to R_a, and the capacitance C is the sum of the diode capacitance, amplifier capacitance, and some distributed stray capacitance. In response to an optical pulse falling on the detector, the load voltage V_{in} will rise and fall exponentially with a time constant RC. In response to a photocurrent I_p, which varies sinusoidally at the angular frequency,

$$\omega = 2\,\pi\,f \tag{6.16}$$

the response of the load voltage will be given by

$$V_{in}\,(f)/I_p\,(f) = R/(1 + j2\,\pi\,f\,C\,R) \tag{6.17}$$

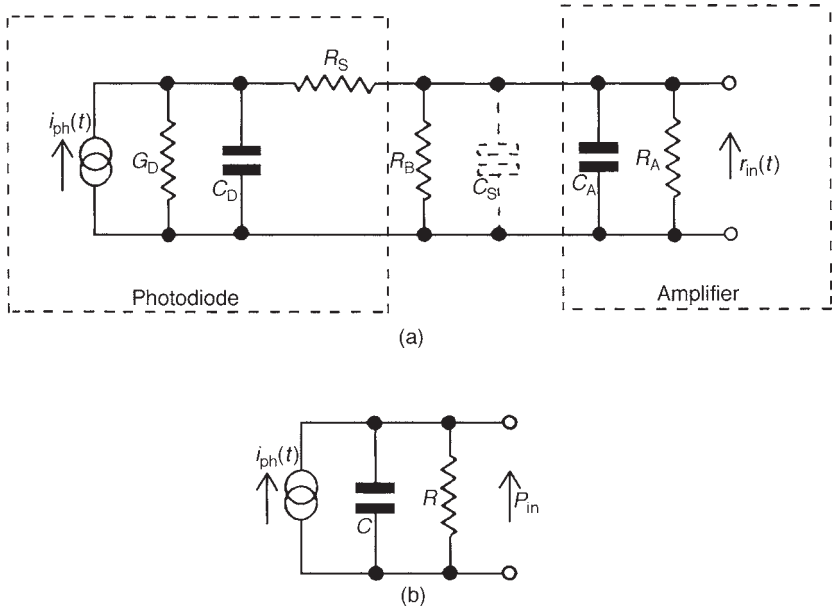


Figure 6.6 Small-signal equivalent circuit for a normally biased photodiode and amplifier: (a) complete cirucuit; (b) reduced circuit obtained by neglecting R_s and lumping together the parallel components.

To obtain good high-frequency response, C must be kept as low as possible; as discussed earlier, the photodiode contribution can normally be kept well below 1 pF. There is ongoing debate concerning the best approach to improve high-frequency response; we must either reduce R or provide high-frequency equalization. As a rule of thumb, no equalization is needed if

$$R < 1/2\ \pi\ C\ \Delta\ f \tag{6.18}$$

where Δ f is the frequency bandwidth of interest. We can also avoid the need for equalization by using a transimpedance feedback amplifier, which is often employed in commercial optical datacom receivers. These apparently simple receiver circuits can exhibit very complex behavior, and it is not always intuitive how to design the optimal detector circuit for a given application. A detailed analysis of receiver response, including the relative noise contributions and tradeoffs between different types of photodiodes, is beyond the scope of this chapter; the interested reader is referred to several good references on this subject [4–13].

6.3. OTHER DETECTORS

The PIN photodiode is definitely the workhorse detector of the fiber optics industry. For direct transceiver connections, they are dependable and cheap. But there are a number of detectors that offer advantages in more complex situations.

Parallel optical interconnects (POIs) require detector arrays, which can be formed from either p-i-n diodes or MSM photoreceivers.

In wavelength-division multiplexing (WDM), the optical signal traveling over one fiber-optic cable is wavelength-separated, coded, and recombined. To properly receive and interpret this combined signal, there is a need for more sensitive detectors, such as avalanche photodiodes, as well as position-sensitive detectors, such as photodiode arrays and MSM photoreceiver arrays, and also wavelength-sensitive detectors, such as resonant-cavity enhanced photodetectors.

The rise of WDM hints at the importance of growing data rates. Schottky barrier photodiodes have always been considered for situations that require faster response than p-i-n diodes (100-Ghz modulation has been reported) but less sensitivity. Metal-semiconductor-metal (MSM) detectors, which are Schottky diode-based planar detectors, are cheap and easy to fabricate, making them desirable for some low-cost applications where there are a number of parallel channels and dense integration. Resonant-cavity enhancement (RECAP) can increase the signal from Schottky barrier and MSM detectors, as well as making high-speed, thin *i* layer, p-i-n diode detectors feasible.

6.3.1. Avalanche Photodiode

PIN diodes can be used to detect light because when photon flux irradiates the junction, the light creates electron-hole pairs with their energy determined by the wavelength of the light. Current will flow if the energy is sufficient to scale the potential created by the PIN junction. This is known as the photovoltaic effect [2, 4–6]. It should be mentioned that the material from which the top layers of a PIN diode is constructed must be transparent (and clean!) to allow passage of light to the junction. If the bias voltage is increased significantly, the photogenerated carriers have enough energy to start an avalanche process, knocking more electrons free from the lattice, which contributes to amplification of the signal. This is known as an avalanche photodiode (APD); it provides higher responsivity, especially in the near-infrared, but it also produces higher noise due to the electron avalanche process. For a PIN photodiode, each absorbed photon ideally creates one electron-hole pair, which, in turn, sets one electron flowing in the external circuit. In this sense, we can loosely compare it to an LED. There is basically a one-to-one relationship between photons and carriers and current. Extending this comparison allows us to say that an avalanche photodiode resembles a laser,

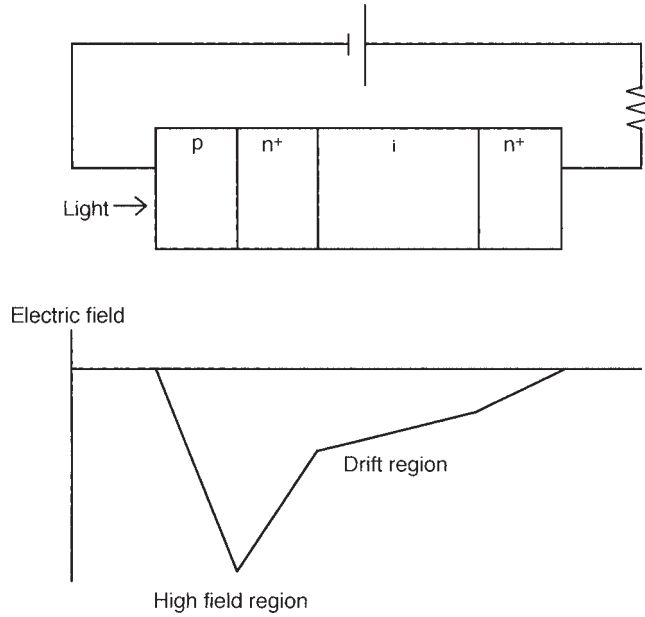


Figure 6.7 Avalanche photodiode.

where the relationship is not one-to-one. In a laser, a few primary carriers result in many emitted photons. In an APD, a few incident photons result in many carriers and appreciable external current.

The structure of the APD, shown in Fig. 6.7, creates a very strong electrical field in a portion of the depletion region. Primary carriers—the free electrons and holes created by absorbed photons—within this field are accelerated by the field, thereby gaining several electron volts of kinetic energy. A collision of these fast carriers with neutral atoms causes the accelerated carrier to use some of its energy to raise a bound electron from the valence band to the conduction band. A free electron-hole pair is thus created. Carriers created in this way, through collision with a primary carrier, are called secondary carriers. This process of creating secondary carriers is known as collision ionization. A primary carrier can create several new secondary carriers, and secondary carriers themselves can accelerate and create new carriers. The whole process is called photomultiplication, which is a form of gain process. The number of electrons set flowing in the external circuit by each absorbed photon depends on the APD's multiplication factor. Typical multiplication ranges in the tens and hundreds. A multiplication factor of 50 means that, on the average, 50 external electrons flow for each photon. The phrase "on the average" is important. The multiplication factor is an average, a

statistical mean. Each primary carrier created by a photon may create more or less secondary carriers and therefore external current. For an APD with a multiplication factor of 50, for example, any given primary carrier may actually create 44 secondary carriers or 53 secondary carriers. This variation is one source of noise that limits the sensitivity of a receiver using an APD.

The multiplication factor M_{DC} of a APD varies with the bias voltage as described in Eq. (6.19):

$$M_{DC} = 1/(1 - V/V_B)^n \tag{6.19}$$

where V_B is the breakdown voltage and n varies between 3 and 6, depending on the semiconductor.

The photocurrent is multiplied by the multiplication factor,

$$I = M\ QE\ q\ \phi\ (\lambda)\ \lambda/h\ c \tag{6.20}$$

as is the responsivity,

$$R(\lambda) = M\ QE\ \lambda\ q/h\ c \tag{6.21}$$

The shot noise in an APD is that of a PIN diode multiplied by M times an excess noise factor, the square root of F, where:

$$F(M) = \beta M + (1 - \beta)(2 - 1/M) \tag{6.22}$$

In this case, β is the ratio of the ionization coefficient of the holes divided by the ionization coefficient of the electrons. In III–V semiconductors F = M. It is also important to remember that the dark current of APDs is also multiplied, according to the same equations as the shot noise (see Section 6.4.2).

Because the accelerating forces must be strong enough to impart energy to the carriers, high bias voltages (several hundred volts in many cases) are required to create the high-field region. At lower voltages, the APD operates like a PIN diode and exhibits no internal gain. The avalanche breakdown voltage of an APD is the voltage at which collision ionization begins. An APD biased above the breakdown point will produce current in the absence of optical power. The voltage itself is sufficient to create carriers and cause collision ionization. The APD is often biased just below the breakdown point, so any optical power will create a fast response and strong output. The tradeoffs are that dark current (the current resulting from generation of electron-hole pairs even in the absence of absorbed photons) increases with bias voltage, and a high-voltage power supply is needed. In addition, as one might expect, the avalanche breakdown process is temperature sensitive, and most APDs will require temperature compensation in datacom applications.

Modern APD design uses separate absorption and multiplication layers (SAMs). This technique allows for high electric fields in the avalanche multiplication region, but lower fields in the low-bandgap-absorbing ternary layer to prohibit tunneling. But the step in the field can cause hole trapping. This is solved

by creating a transition region consisting of one or more intermediate steps, known as the separate absorption grading and multiplication (SAGM) structure. In order to make it easier to fabricate SAGM APDs, and at the same time to avoid premature edge breakdown, a partial charge sheet layer is introduced between the multiplication and the grading region. This structure is known as the separate absorption, grading, charge sheet, and multiplication (SAGCM) structure. It is fairly difficult for available SAGM and SAGCM APDs to respond reliably to 10 GB/s range signals. This is due to the speed limitations inherent in the avalanche process, the electron and hole ionization rates.

Superlattice avalanche photodiodes (SL-APDs) use the bandgap discontinuity to enhance the ionization rate and can overcome the limitation on conventional APD performance. This in turn leads to a multiple quantum well (MQW) structure. A gain bandwidth product of 150 Ghz, with a dark current of 6 nA, has been reported for a layer structure with a 0.23-μm-thick SL multiplication layer with 8-nm-thick InAlGaAs wells (Eg = 1 eV) and 13-μm-thick InAlAs barriers, a p-InP field buffer layer (~52 nm), and a 1-μm-thick p InGaAs absorption layer. (7)

6.3.2. Photodiode Array

Photodiode arrays are beginning to be used as detectors in parallel optics for supercomputers. They are also the detector used for spectrometers (along with charge coupled detector (CCD) arrays), which makes them useful to test fiber optics systems. Photodiode arrays are also being considered for wavelength-division multiplexing (WDM).

In Section 6.2 we discussed the physics guiding photodiode operation. In a photodiode array, the individual diode elements respond to incident flux by producing photocurrents, which charge individual storage capacitors. Indium gallium arsenide (InGaAs) photodiode arrays are the materials used in spectroscopic applications. They can be embedded in erbium doped fiber amplifiers (EDFAs). Crosstalk, or signal leakage between neighboring pixels, is normally a concern in array systems. In InGaAs photodiode arrays it is limited to nearest-neighbor interactions. However, the EDFAs do exhibit crosstalk, so there is some discussion of using different amplifiers, such as transimpedance amplifiers (TIA arrays). In the not too distant future, this type of detector will operate at 10 Gbit/s [15].

While some wavelength division multiplexers (WDM) filter out the wavelength dependent signal, another common way to separate the signal is by using a grating to place different wavelengths at different physical positions, as is done in spectrometers. In dense WDM, as many as 100 optical channels can be used. Photodiode arrays are an obvious detector choice. In addition to detecting the signal, they offer performance feedback to the tunable lasers. At this time, photodiode arrays have not been used in WDM, but their use would be similar to spectrometer design.

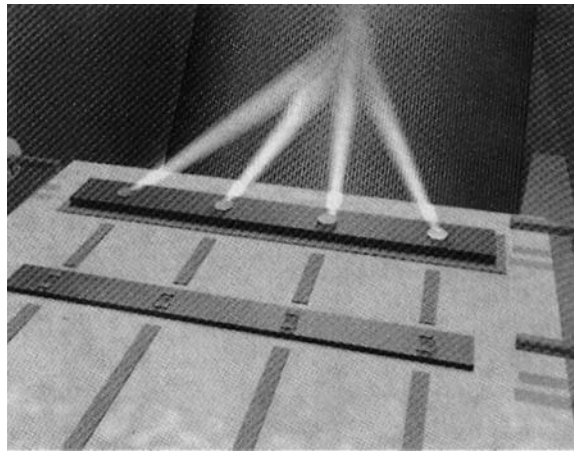


Figure 6.8 An optical demultiplexing receiver array, where data channels are focused onto an array of high-speed InGaAs photodiodes. The signals are then amplified using a trans-impedance amplifier (TIA) array [15].

6.3.3. Schottky Barrier Photodiodes

A variation from the PIN diode structure is shown in Fig. 6.9. This is known as a Schottky barrier diode; the top layer of semiconductor material has been eliminated in favor of a reverse-biased, metal-semiconductor-metal (MSM) contact. The metal layer must be thin enough to be transparent to incident light, about 10 nm; alternate structures using interdigital metal transducers are also possible. The advantage of this approach is improved quantum efficiency, because there is no recombination of carriers in the surface layer before they can diffuse to either the ohmic contacts or the depletion region.

When you want to make a PIN detector respond faster, your only choice is to make the inductor layer thinner. This decreases the signal. Some architectures have been designed to try to get around this problem, such as resonant-cavity photodetectors (see Section 6.3.5), or placing a number of PIN junctions into an optical waveguide, known as traveling wave photodetectors. But Schottky barrier photodiodes are another traditional alternative.

Schottky barrier photodiodes, unlike the prior detectors discussed, are not p-i-n based. Instead, a thin metal layer replaces half of the p-n junction. However, it does result in the same voltage characteristics: when an incident photon hits the metal layer, carriers are created in the depletion region, and their movement again sets current flowing. The voltage and field relations derived for PIN junctions can be applied to Schottky barrier diodes by treating the metal layer as if it were an extremely heavily doped semiconductor.

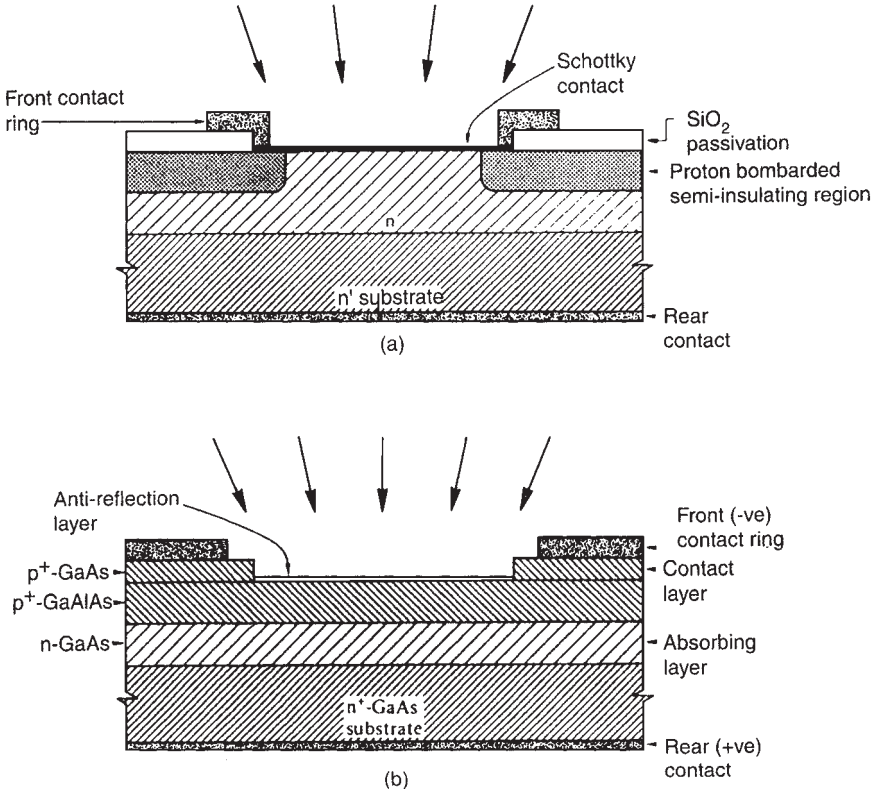


Figure 6.9 Schottky barrier diode.

The metal layer is a good conductor, and so electrons leave the junction immediately. This results in faster operation—up to 100-Ghz modulation has been reported [14]. It also lowers the chance of recombination, improving efficiency, and increases the types of semiconductors that can be used, since one is less limited by lattice-matching fabrication constraints. However, their sensitivity is lower than PIN diodes with the usual sized intrinsic layer. Resonant enhancement (see Section 6.3.5) has been experimented with in order to increase sensitivity.

The main problem in designing Schottky barrier photodiodes for fiber-optic wavelengths is that suitable metals for use on InGaAs substrates are not available. Usually Schottky barrier diodes use a thin gold layer, with a current collecting low-resistance thick ring of gold the edge of the active region. But to avoid low Schottky barrier height for InGaAs, a thin top layer of AlInAs is used instead. Intermediate bandgap material AlGaInAs is sandwiched in-between InGaAs and AlInAs layers to prevent band discontinuity [7].

Figure 6.10 A diagram of an MSM detector. [MSM Photodetectors by Pekka Kuivalainen, Nina Hakkarainen, Katri Honkanen, Helsinki University of Technology Electron Physics Laboratory].

Single Schottky barrier photodiodes are not currently used in fiber-optic applications, but they are being considered as communication speeds increase.

6.3.4. Metal-Semiconductor-Metal (MSM) Detectors

A MSM photodetector is made by forming two Schottky diodes on top of a semiconductor layer. The top metal layer has two contact pads and a series of "interdigitated fingers" (Fig. 6.10), which form the active area of the device. One diode gets forward biased, and the other diode gets reverse biased. When illuminated, these detectors create a time-varying electrical signal. MSM photoreceiver arrays have been used as parallel optical interconnects. They are also being considered for use in smart pixels.

Like Schottky diodes, speed is an advantage of MSM detectors. The response speed of MSM detectors can be increased by reducing the finger spacing, and thus the carrier transit times. However, for very small finger spacing, the thickness of the absorption layer becomes comparable to the finger spacing and limits the high-speed performance.

The other advantages of MSM detectors are their ease in fabrication and integration into electrical systems. The MSM photodiodes do not require doping, and they are essentially identical to the gate metallization of field-effect transistors (FETs) [7]. They are being placed in systems using arrays and smart pixels.

MSM detectors have the same disadvantages of other Schottky-based systems—weak signal and noise. They have some advantages in systems with a number of parallel channels and dense integration of detectors, which could eventually be applied to WDM systems.

As with Schottky barrier photodiodes, experiments have been done to increase their signal using resonant enhancement (see Section 6.3.5).

For a MSM detector deposited on a photoconductive semiconductor with a distance L between the electrodes, the photocurrent and time-varying resistance can be calculated the same as for any photoconductive detector [14]. Let's assume the detector is irradiated by a input optical flux P, at a photon energy hv.

$$I_{ph} = GP\ QE\ q\ \phi\ (\lambda)\ \lambda/h\ c \tag{6.23}$$

where QE is the quantum efficiency and G is the photoconductive gain, which is the ratio of the carrier lifetime to the carrier transit time. It can be seen that increasing the carrier lifetime decreases the speed, but increases the sensitivity, which is what you would expect.

The time-varying resistance R(t), is dependent on the photoinduced carrier density $N(t)$.

$$N(t) = QE\ \phi\ (\lambda)\ \lambda/h\ c \tag{6.24}$$

The time-varying resistance is

$$R(t) = L/(eN(t)\mu wd) \tag{6.25}$$

where μ is the sum of the electron and whole mobilites, w is the length along the electrodes excited by the light, and d is the effective absorption depth into the semiconductor.

On silicon and GaAs photodiodes, the transparent metal layer is made from indium tin oxide (ITO) or cadmium tin oxide (CTO). CTO can also be used for InGaAs photodetectors, but ITO is not appropriate due to the large absorption at the 1.0–1.6 μm wavelength region. Other variations using gold and SiN_x antireflection coatings have also been fabricated [7].

6.3.5. Resonant-Cavity Enhanced Photodetectors

The principle of resonant-cavity enhanced photodetectors (RECAPs) is that you put a fast photodetector into a Fabry-Perot (FP) cavity to enhance the signal

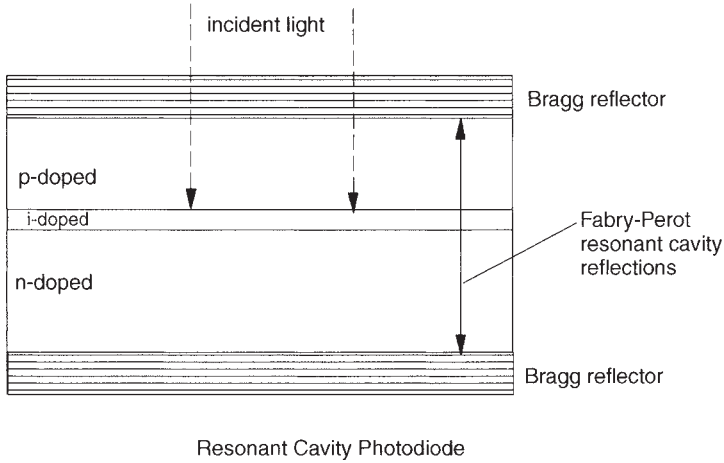


Figure 6.11 The diagram of a resonant-cavity photodetector. Note how the PIN diode is located between the Bragg Reflector Stacks, creating a Fabry-Perot resonant cavity.

magnitude. In practice, the Fabry-Perot cavity's end mirrors are made of quarter-wave stacks of GaAs/AlAs, InGaAs/InAlAs, or any other semiconductor or insulator materials that have the desired refractive index contrast. A typical example of a fast photodetector is a PIN photodetector with a very thin i-layer. By making the inductor layer thinner, you increase the speed of the detector but decrease the signal. Avalanche photodiodes (6.3.1), Schottky barrier photodiodes (see 6.3.3), and MSM detectors (see Section 6.3.4) can also have their signal magnitude increased using RECAP techniques.

By placing the detector in an FP cavity, you pass the light through it several times, which results in an increase of signal. However, due to the resonance properties of the FP cavity, this detector is highly wavelength selective. It is also highly sensitive to the position of the detector in the FP cavity, since it is a resonance effect.

The maximum quantum efficiency QE_{max} of a RECAP detector can be calculated by

$$QE_{max} = \left\{ \frac{(1 + R_2 e^{-\alpha d})(1 - R_1)}{(1 - \sqrt{R_1 R_2}\, e^{-ad})^2} \right\} * (1 - e^{-ad}) \tag{6.26}$$

where R_1 and R_2 are the reflectivity of the top and bottom mirrors, α is the absorption of the active area of the detector, and d is the thickness of the active area of the detector [16]. If the reflection of the bottom mirror is $R_2 = 0$, this reduces to the quantum efficiency of a photodiode, only with an ideal total transmission through the top electrode, as in Eq. (6.2),

$$QE = (1 - R)T(1 - e^{-\alpha W}) \tag{6.2}$$

The resonant-cavity enhancement, RE, can allow RCE detectors to have quantum efficiencies of nearly unity at their peak wavelength.

$$RE = \left\{ \frac{(1 + R_2 e^{-ad})}{(1 - \sqrt{R_1 R_2}\, e^{-ad})^2} \right\} \tag{6.27}$$

Because WDM applications are by their very nature wavelength dependent, the wavelength selectivity of RECAP may turn out be a useful design feature. RECAP is not yet being used in the field, but reported experimental quantum efficiencies have reached 82% for PIN diodes, and 50% improvements of photocurrent for Schottky barrier photodiodes [16].

6.4. NOISE

Any optical detection or communication system is subject to various types of noise; there can be noise in the signal, noise created by the detector, and noise in the electronics. A complete discussion of noise sources has already filled several good reference books; since this is a chapter on detectors, we will briefly discuss

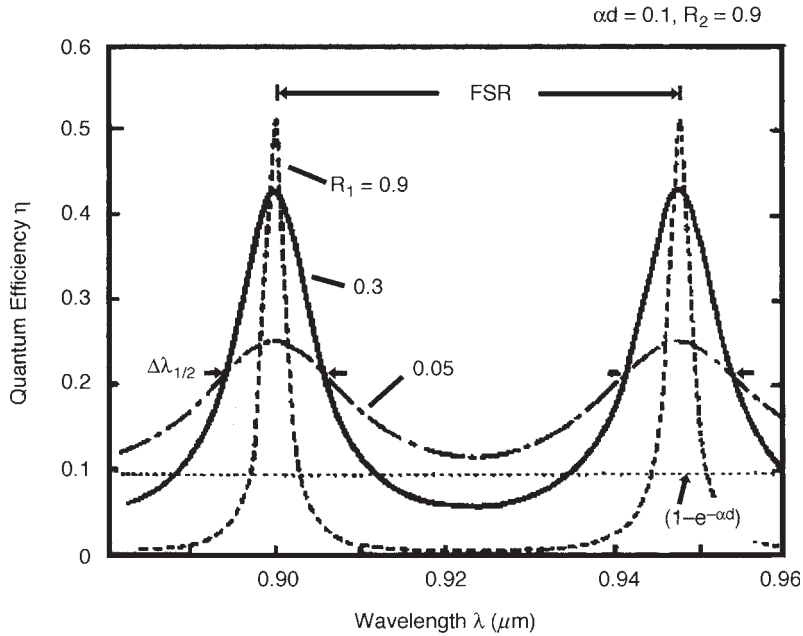


Figure 6.12 The quantum efficiency of resonant cavity detectors as a function of wavelength [16].

noise created by detectors. For a more complete discussion, the reader is referred to treatments by Dereniak and Crowe [6], who have categorized the major noise sources. The purpose of the detector is to create an electrical current in response to incident photons. It must accept highly attenuated optical energy and produce an electrical current. This current is usually feeble because of the low levels of optical power involved, often only in the order of nanowatts. Subsequent stages of the receiver amplify and possibly reshape the signal from the detector. Noise is a serious problem that limits the detector's performance. Broadly speaking, noise is any electrical or optical energy apart from the signal itself. Although noise can and does occur in every part of a communication system, it is of greatest concern in the receiver input. The reason is that the receiver works with very weak signals that have been attenuated during transmission. Although very small compared to the signal levels in most circuits, the noise level is significant in relation to the weak detected signals. The same noise level in a transmitter is usually insignificant because signal levels are very strong in comparison. Indeed, the very limit of the diode's sensitivity is the noise. An optical signal that is too weak cannot be distinguished from the noise. To detect such a signal, we must either reduce the noise level or increase the power level of the signal. In the following

sections, we will describe in detail several different noise sources. In practice, it is often assumed that the noise in a detection system has a constant frequency spectrum over the measurement range of interest. This is so-called white noise or Gaussian noise and is often a combination of the effects we will describe here.

6.4.1. Noise and Amplification

The amplification stages of the receiver amplify both the signal and the noise. Some AC techniques, such as lock-in amplification, can help separate the signal from the noise. To illustrate the principle, we will digress briefly to describe the example of the lock-in amp, even though it is not commonly used in datacom systems. Lock-in amplification, or chopping, is a technique used to limit noise, although it can also be used to detect dc signals that are deliberately encoded with a known modulation. A mechanical light chopper (which looks somewhat like a fan) is placed between the signal and the detector. The amplifier is connected to the chopper, and so the amplifier "knows" from this reference signal when the signal to be detected is on and when it is off. As you can imagine, this makes for very precise background subtraction. However, that is not all the lock-in amplifier does.

The secret of the lock-in amplifier is that it narrows the bandwidth of the detector, and the narrower that bandwidth, the more precise the measurement. It does this by creating a weak periodic signal, and low pass filtering it (or narrowing the bandwidth). This helps eliminate 1/f noise (flicker noise, drifts, etc.). The signal you have after lock-in amplification is a differential signal, not a linear signal. It is good for determining changes in signal, not signal magnitude. Horowitz and Hill discuss lock-in amplification in [17]:

> In order to illustrate the power of lock-in detection, we usually set up a small demonstration for our students. We use a lock-in to modulate a small LED of the kind used for panel indicators, with a modulation rate of a kilohertz or so. The current is very low, and you can hardly see the LED glowing in normal room light. Six feet away a phototransistor looks in the general direction of the LED, with its output fed to the lock-in. With the room lights out, there's a tiny signal from the phototransistor at the modulating frequency (mixed with plenty of noise), and the lock-in easily detects it, using a time constant of a few seconds. Then we turn the room lights on (fluorescent), at which point the signal from the phototransistor becomes just a huge messy 120 Hz waveform, jumping in amplitude by 50 db or more. The situation looks hopeless on the oscilloscope, but the lock-in just sits there, unperturbed, calmly detecting the same LED signal at the same level. You can check that it's really working by sticking your hand in between the LED and detector. It's darned impressive. (p. 631)

Similar tricks to narrow the bandwidth are used in signal averaging, boxcar integration, multichannel scaling, pulse height analysis, and phase-sensitive detection. For example, boxcar averaging takes its name from gating the signal detection time into a repetitive train of N pulse intervals, or "boxes," during which the signal is present. Since noise that would have been accumulated during times when the gating is off is eliminated, this process improves the signal-to-noise ratio by a factor of the square root of N for white, Johnson, or shot noise. (This is because the integrated signal contribution increases as N, while the noise contribution increases only as the square root of [6].) Narrowing the bandwidth of a fiber-optic receiver can also have beneficial effects in controlling RIN and modal noise (see Chapter 7). The use of differential signaling is also common in datacom receiver circuits, although they typically do not use lock-in amps but rather solid-state electronics such as operational amplifiers (see Chapter 16).

6.4.2. Shot Noise

Shot noise occurs in all types of radiation detectors. It is due to the quantum nature of photoelectrons. Since individual photoelectrons are created by absorbed photons at random intervals, the resulting signal has some variation with time. The variation of detector current with time appears as noise; this can be due to either the desired signal photons or background flux (in the latter case, the detector is said to operate in a Background Limited in Performance, or BLIP, mode). To study the shot noise in a photodiode, we will consider the photodetection process. An optical signal and background radiation are absorbed by the photodiode, whereby electron-hole pairs are generated. These electrons and holes are then separated by the electric field and drift toward the opposite sides of the p-n junction. In the process, a displacement current is induced in the external load resister. The photocurrent generated by the optical signal is I_p. The current generated by the background radiation is I_b. The current generated by the thermal generation of electron-hole pairs under completely dark environment is I_d. Because of the randomness of the generation of all these currents, they contribute shot noise given by a mean square current variation of

$$I_s^2 = 2q(I_p + I_b + I_d)\, B \tag{6.28}$$

where q is the charge of an electron (1.6×10^{-19} coulomb) and B is the bandwidth. The equation shows that shot noise increases with current and with bandwidth. Shot noise is at its minimum when only dark current exists, and it increases with the current resulting from optical input.

6.4.3. Thermal Noise

Also known as Johnson or Nyquist noise, thermal noise is caused by randomness in carrier generation and recombination due to thermal excitation in a conductor. It results in fluctuations in the detector's internal resistance, or in any

resistance in series with the detector. These resistances consist of R_j, the junction resistance; R_s, the series resistance; R_l, the load resistance; and R_i, the input resistance of the following amplifier. All the resistances contribute additional thermal noise to the system. The series resistance R_s is usually much smaller than the other resistance and can be neglected. The thermal noise is given by

$$I_t^2 = 4kTB((1/R_j) + (1/R_l) + (1/R_i)) \tag{6.29}$$

where k is the Boltzmann constant (1.38×10^{-23} J/K), T is absolute temperature (kelvin scale), and B is the bandwidth.

6.4.4. Other Noise Sources

Generation recombination and 1/f noise are particular to photoconductors. Absorbed photons can produce both positive and negative charge carriers, some of which may recombine before being collected. Generation recombination noise is due to the randomness in the creation and cancellation of individual charge carriers. It can be shown [4] that the magnitude of this noise is given by

$$I_{gr} = 2\ I\ (\tau\ B/N\ (1 + (2\ \pi\ f\ \tau)^2))^{½} = 2\ q\ G\ (\varepsilon\ E\ A\ B)^{½} \tag{6.30}$$

where I is the average current due to all sources of carriers (not just photocarriers), τ is the carrier lifetime, N is the total number of free carriers, f is the frequency at which the noise is measured, G is the photoconductive gain (number of electrons generated per photogenerated electron), E is the photon irradiance, and A is the detector active area.

Flicker, or so-called 1/f noise, is particular to biased conductors. Its cause is not well understood, but it is thought to be connected to the imperfect conductive contact at detector electrodes. It can be measured to follow a curve of $1/f^{\beta}$, where β is a constant that varies between 0.8 and 1.2; the rapid falloff with 1/f gives rise to the name. Lack of good ohmic contact increases this noise, but it is not known if any particular type of electrical contact will eliminate this noise. The empirical expression for the noise current is

$$I_f = \alpha\ (i\ B/f^{\beta})^{½} \tag{6.31}$$

where α is a proportionality constant, i is the current through the detector, and the exponents are empirically estimated to be $\alpha = 2$ and $\beta = 1$. Note that this is only an empirical expression for a poorly understood phenomenon; the noise current does not become infinite as *f* approaches zero (DC operation).

There may be other noise sources in the detector circuitry as well; these can also be modeled as equivalent currents. The noise sources described here are uncorrelated and thus must be summed as rms values rather than a linear summation (put another way, they add in quadrature or use vector addition), so that the total noise is given by

$$I_{tot}^2 = I_f^2 + I_{gr}^2 + I_s^2 + I_t^2 \tag{6.32}$$

Usually, one of the components in the above expression will be the dominant noise source in a given application. When designing a data link, one must keep in mind likely sources of noise, their expected contribution, and how to best reduce them. Choose a detector so that the signal will be significantly larger than the detector's expected noise. In a laboratory environment, cooling some detectors will minimize the dark current; this is not practical in most applications. Other rules of thumb can be applied to specific detectors as well. For example, although Johnson noise cannot be eliminated, it can be minimized. In photodiodes, the shot noise is approximately 3 X greater than Johnson noise if the dc voltage generated through a transimpedance amplifier is more than about 500 mV. This results in a higher degree of linearity in the measurements while minimizing thermal noise. The same rule of thumb applies to PbS- and PbSe-based detectors, though care should be taken not to exceed the maximum bias voltage of these devices, or catastrophic breakdown will occur. For the measured values of NEP, detectivity, and specific detectivity to be meaningful, the detector should be operating in a high-impedance mode so that the principal source of noise is the shot noise associated with the dark current and the signal current. Although it is possible to use electronic circuits to filter out some types of noise, it is better to have the signal much stronger than the noise by either having a strong signal level or a low noise level. Several types of noise are associated with the photodiode itself and with the receiver; for example, we have already mentioned multiplication noise in an APD, which arises because multiplication varies around a statistical mean.

6.4.5. Signal-to-noise Ratio

Signal-to-noise ratio (SNR) is a parameter of describing the quality of signals in a system. SNR is simply the ratio of the average signal power, S, to the average noise power, N, from all noise sources.

$$\mathrm{SNR} = \mathrm{S/N} \tag{6.33}$$

SNR can also be written in decibels as

$$\mathrm{SNR} = 10 \log_{10} (\mathrm{S/N}) \tag{6.34}$$

If the signal power is 20 mW and the noise power is 20 nW, the SNR ratio is 1000, or 30 dB. A large SNR means that the signal is much larger than the noise. The signal power depends on the power of the incoming optical power. The specification for SNR is dependent on the application requirements.

For digital systems, bit error rate (BER) usually replaces SNR as a performance indicator of system quality. BER is the ratio of incorrectly transmitted bits to correctly transmitted bits. A ratio of 10^{-10} means that one wrong bit is received for every 10 billion bits transmitted. Similarly with SNR, the specification for

BER is also dependent on the application requirements. SNR and BER are related. In an ideal system, a better SNR should also have a better BER. However, BER also depends on data-encoding formats and receiver designs. There are techniques to detect and correct bit errors. We cannot easily calculate the BER from the SNR, because the relationship depends on several factors, including circuit design and bit error correction techniques. For a more complete overview of BER and other sources of error in a fiber-optic data link, see Chapter 7.

ACKNOWLEDGMENT

The contributions of Bill Reysen, AMP Incorporated, Lytel Division, to the preparation of this chapter are gratefully acknowledged.

REFERENCES

1. DeCusatis, C. J Sher. 1997. Fundamentals of detectors. In *Handbook of applied photometry*, ed. C. DeCusatis, pp. 101–132. New York: AIP Press.
2. Burns, G. 1985. *Solid State Physics*. Orlando, Fl.: Academic Press, pp. 323–334.
3. NBS Special Publication SP250-17. The NBS Photodetector Spectral Response Calibration Transfer Program.
4. Lee T. P., and T. Li. 1979. Photodetectors. In *Optical fiber communications*, ed. S. E. Miller and A. G. Chynoweth. New York: Academic Press.
5. Gowar, J. 1984. *Optical communication systems*. Englewood Cliffs, N.J.: Prentice Hall.
6. Dereniak, E., and D. Crowe. 1984. *Optical radiation detectors*. New York: John Wiley & Sons.
7. Bandyopadhyay, A., and M. Jamal Deen. 2001. Photodectors for optical fiber communications. In *Photodetectors and fiber optics*, H.S. Nalwa (ed.). New York: Academic Press.
8. Graeme, J. 1994, June 27. Divide and conquer noise in photodiode amplifiers. *Electronic Design*, pp. 10–26.
9. Graeme, J. 1994, November 7. Filtering cuts noise in photodiode amplifiers. *Electronic Design*, pp. 9–22.
10. Graeme, J. 1987, October 29. FET op amps convert photodiode outputs to usable signals. *EDN*, p. 205.
11. Graeme, J. 1982, May 7. Phase compensation optimizes photodiode bandwidth. *EDN*, p. 177.
12. Bell, D. A. 1985. *Noise and the solid state*. New York: John Wiley & Sons.
13. Burt, R., and R. Stitt. 1988, September 1. Circuit lowers photodiode amplifer noise. *EDN*, p. 203.
14. Garmire, Elsa. 2001. *Sources, modulators and detectors for fiber optic communication systems. Handbook of Optics* IV. New York: McGraw-Hill.
15. Cohen, Marshall J. 2000, August. Photodiode arrays help meet demand for WDM. *Optoelectronics World, Supplement to Laser Focus World*.
16. Selim Ünlü, M., and Samuel Strite. 1995. Resonant cavity enhanced (RCE) photonic devices" http://photon.bu.edu/selim/papers/apr-95/node1.html.
17. Horowitz, P., and W. Hill. 1980. *The art of electronics*. Cambridge: Cambridge University Press, ch. 14, pp. 628–631.

This page intentionally left blank

7

Receiver Logic and Drive Circuitry

Ray D. Sundstrom
Motorola, Incorporated, Chandlec, Arizona

Eric Maass
Motorola, Incorporated, Tempe, Arizona

Linear/Limiting Receiver Contribution by Scott Kipp, Brocade Corporation

7.1. SYSTEM OVERVIEW

In data communications applications, the data and address information is generally supplied through parallel lines. The parallel information can be directly converted, line for line, into parallel optical signals: alternatively, the parallel information can first be converted into a serial bit stream and then transmitted through a serial optical link. The parallel optical approach is illustrated through a block diagram in Fig. 7.1a, and the serial optical approach is illustrated in Fig. 7.1b. In the parallel optoelectrical interface, illustrated in Fig. 7.1a, each parallel electrical line has a laser driver and laser associated with it. The light is sent through parallel channels, such as fiber ribbon composed of several optical fibers. The light from each channel is guided to a photodetector, and the signal from each photodetector is amplified through a transimpedance and postamplifier, one for each channel. The array of postamplifiers, including conversion to appropriate digital signal levels, provides a parallel electrical output similar to the parallel electrical lines that were the original inputs. In the serial optical-electrical interface, illustrated in Fig. 7.1b, the parallel electrical lines are multiplexed to provide a serial signal output. The parallel to serial conversion also requires clock generation. The serialized electrical signal is provided through a laser or light-emitting diode (LED) driver circuit to the laser or LED. The driver circuit may involve feedback circuitry from the laser or LED to keep the drive current at an appropriate level for the laser or LED operation. The resulting serial optical signal is

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

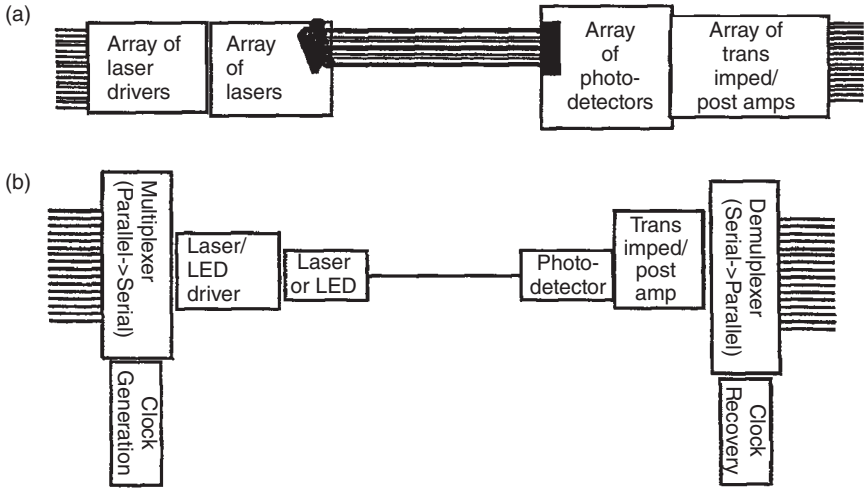


Figure 7.1 (a) Block diagram for parallel optoelectrical interface, involving arrays of optoelectronic interfaces. (b) Block diagram for serial optoelectrical interface, involving parallel–serial–parallel conversions.

coupled to optical fiber or other light-transmitting medium. On the receive side, the serial optical signal is coupled to a single photodetector. The signal from the photodetector is generally a very small current, which is converted into a voltage through the transimpedance amplifier, and amplified with a post- or limiting amplifier. The resulting serial electrical signal must then be converted to a parallel signal through a demultiplexer. This serial-parallel conversion requires clock-recovery circuitry. The clock-recovery circuitry involves special requirements, such as recovering clock signals from a variety of bit patterns; these special requirements and the design considerations involved are described later in this chapter and elaborated on further in Chapter 16.

The circuitry associated with these systems can be implemented in various semiconductor technologies, including GaAs, silicon bipolar, complementary metal oxide semiconductor (CMOS), and BiCMOS. GaAs has a particular advantage for the transimpedance function on the receive side in that the photodiode can readily be integrated with this function. The high mobility associated with GaAs facilitates meeting the high-frequency requirements of any circuit in the system. The low transconductance of GaAs field-effect transistors (FETs) is a disadvantage for the postamplifier function, generally requiring more amplifier stages. A major disadvantage of GaAs historically has been the high cost relative to silicon solutions. Bipolar technology has advantages in high bandwidth and high transconductance, which is particularly helpful for the transimpedance func-

tion, and high voltage gain, which is helpful for the postamplifier function. The high-frequency performance of bipolar circuits using emitter-coupled logic (ECL) design approaches is useful for the multiplexer and demultiplexer functions.

The high-static power consumption associated with high-speed ECL is not an issue if data are expected to be continually transmitted and received. Bipolar integrated circuits tend to cost less than similar GaAs-integrated circuits but more than CMOS circuits. Small geometry CMOS could be used for several, if not all, functions and would be expected to provide the lowest cost solution. The relatively low transconductance of MOS transistors is a disadvantage for both the transimpedance amplifier and the postamplifier function. CMOS would also encounter disadvantages for high-speed, quadrature clock signals for the clock-generation and clock-recovery functions. Using differential techniques with CMOS for the multiplexer and demultiplexer functions has some advantages. BiCMOS has an advantage at higher levels of integration because it has the flexibility associated with both bipolar and MOS transistors. The main disadvantage of BiCMOS is the relatively high cost and long cycle time due to the process complexity.

For LEDs, the low current is usually 0 mA, meaning that the diode is completely shut off and emits no light when it is low. The high current is set by the system designer and depends on the system requirements. The higher the current, the higher the light output of the LED. Of course, there is a maximum current limit for every device (beyond which the part's output or reliability will begin to degrade). A normal "on" current for LEDs is in the range of 20–100 mA.

For lasers, the low current can also be 0, but it is usually some value large enough to keep the laser active. The advantage of keeping the laser active is that it reduces the turn-on delay of the device. Depending on device size and system requirements, the laser may switch from low = 1 mA to high = 5 mA, or as great as low = 100 mA and high = 300 mA. A typical receiver would include a p-type-intrinsic-n-type (PIN) diode or an avalanche photodiode (APD) to receive the photons and convert them into a proportional current. This current is fed into a transimpedance amplifier (TZA), which converts the current into a voltage.

7.2. ELECTRICAL INTERFACE

Transistor-transistor logic (TTL) and CMOS are very common interfaces that are easy to use and do not require termination if the technologies are not pushed to the limit of their bandwidth capabilities. For higher performance, TTL and CMOS can be transmitted over controlled impedances and terminated correctly to reduce problems with noise and reflections. The TTL interface can be used up to approximately 100 MHz, and low-voltage CMOS can be used up to approximately 250 MHz. However, the rail-to-rail voltage swings of these two logic families can consume much power and generate much noise at such high frequen-

cies. Some CMOS designs now include low-voltage differential interfaces to alleviate the noise and power problems. Differential ECL and Current-Mode Logic interfaces are better suited for very high performance electrical interconnect. These two logic configurations were designed to drive controlled impedance transmission lines and have very good noise immunity due to the differential operation. These two interfaces also have smaller voltage swings, in the range from 0.25 to 1.0 V, that can consume less power at high frequencies. The lower voltage swing also reduces electromagnetic interference. These logic interfaces have been used at data rates as high as 10 Gb/s.

7.3. OPTICAL INTERFACE

The optical devices used in the transmitter can have a large effect on system performance. All devices have parasitic properties that the circuit designer must take into account when designing an optical link. Light output of LEDs and lasers varies with temperature and age. Photodiode efficiency is affected by the reverse-bias voltage. Parasitics of the device packages and circuit board can play a major factor in system performance at high frequencies. Higher performance systems usually require that the driver and laser be packaged in a module to reduce lead inductance. The photodiode and TZA also need to be packaged in a module for the same reason.

On the receiver side, noise current at the input of the TZA is crucial in determining how small a signal the amplifier can reliably receive. It is a common practice to specify noise at the input of an amplifier. Input noise is determined by adding the contribution of each noise source in a circuit and calculating an equivalent value as if it were at the input. The noise must be about an order of magnitude less than the minimum input signal expected to be received. This ratio may vary, dependent on the level of errors that can be tolerated. There is always a compromise between bandwidth and input noise current when designing a TZA. The parasitic capacitance of the PIN diode is also one of the dominant factors in determining the bandwidth of the receiver. The diode capacitance and input impedence of the TZA often form the dominant pole in the receiver.

All types of laser devices have similar design challenges. Lasers have a long turn-on delay if shut off completely. To minimize this turn-on delay, lasers are generally prebiased with a DC bias current and modulated with a switching current. This configuration greatly enhances optical performance but complicates the design of the transmitter and receiver: There are two separate currents that need precise control for optimal performance. The objective is to bias the DC current precisely at the threshold current (where the laser begins to emit light). This allows the maximum difference between a high level and a low level for the least amount of power. Maintaining proper biasing of the laser is not a trivial task.

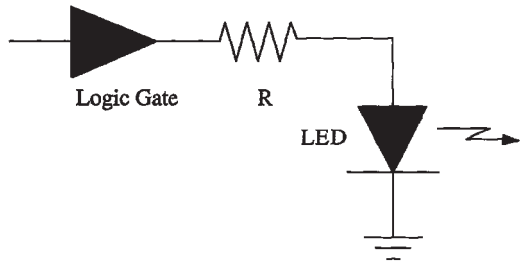


Figure 7.2 Diagram of simple output.

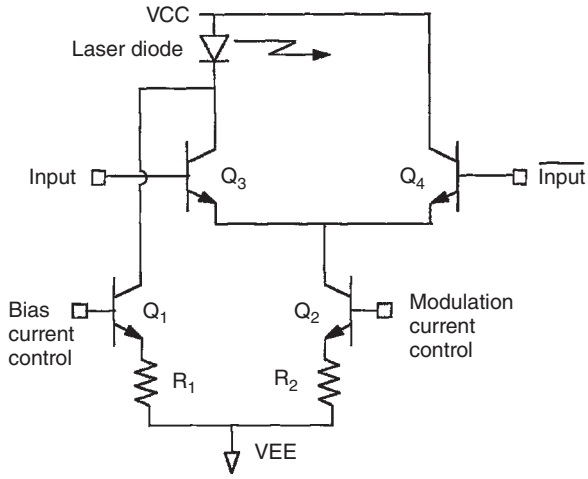


Figure 7.3 Simplified diagram of laser output driver.

7.3.1. Driver Circuitry

The optical driver can vary from a simple switch to a complicated feedback system accounting for many variables. The simple driver may consist of a CMOS or TTL output tied to a resistor and a LED or laser diode, as shown in Fig. 7.2. If system performance is not pushed to the limits, this might be the simplest and therefore perhaps the best solution. Lasers can be switched on and off, but their performance increases significantly if they are not shut off completely. The performance can be improved by providing both a DC bias current and a switching or modulation current. A differential gate can provide a controlled current source for modulation and another controlled current source continuously passed through the laser diode for the bias current. Bias current is a constant current flowing

through the laser. The bias current should be set at the threshold current—the point at which the laser just starts to emit light, as illustrated in Fig. 7.3. If the laser is biased below the threshold current, the turn-on delay will increase. If it is biased above the threshold current, the difference between the light emission corresponding to a 0 and a 1 will be decreased. It is important to maximize the difference between the light emission corresponding to a 0 and a 1 because this is one of the factors that determines how far the link can transmit—the larger the difference in the light signals for a 0 and a 1, the larger the signal will be at the receiver.

Figure 7.3 is a transistor-level diagram of an output driver that implements the bias and modulation currents for the laser. A bias control voltage is forced on the base of *Q1*, which puts a voltage across *R1*. The voltage across *R1* is the bias control voltage minus the base-emitter voltage of *Ql*. The current in the emitter of *Ql* is the voltage across *R1* divided by the resistance of *R1*. This current minus the base current is the current in the collector of *Q1*, which will always flow through the laser diode. This is one method of obtaining the laser bias current. The bias control voltage can be either fixed or controlled by some other monitor circuit. The modulation current is controlled in a similar fashion by the modulation current control voltage. The modulation current, however, is not always directed through the laser diode. Q3 and Q4 make a differential gate that determines whether the current flows through the laser diode. If the voltage on the base of Q3 is higher than the voltage on the base of Q4, the current will flow through the laser diode, through Q3, and into the current source created by Q2 and *R2*. Because the modulation current is flowing through the laser, the optical output is increased corresponding to an optical high level. Alternatively, if the voltage on the base of Q4 is high and that on Q3 is low, the current will flow from the power supply through Q4 and into *Q2*. Because the current is not flowing through the laser, the optical output is decreased, corresponding to an optical low state. As data rates and transmission distances increase, the output of the laser diode needs to be more tightly controlled. Because the laser output power can vary with temperature and lifetime, some method of monitoring the light output and for feedback to the driver is generally incorporated in higher performance systems. With edge-emitting lasers that emit light from both the front and the back facets, a photodetector or monitor diode is commonly mounted on the back side of the laser. The signal from the monitor diode determines the bias and modulation current adjustments in a closed-loop feedback system.

7.3.2. Receiver Circuitry

The optical receiver design can be virtually an art. The challenge involves tradeoffs in improving some parameters at the expense of other parameters. The design is optimized by appropriate compromise *so* that the receiver has the best

combination of high sensitivity, correct bandwidth, low noise, large dynamic range, and low power. Until direct processing of optical information becomes a technology, optical information must be converted to electrical information for processing. The conversion commonly uses a PIN diode. The intrinsic layer between the p-type and n-type layers allows the formation of a very large depletion region when a reverse DC bias is applied. Photons with sufficient energy are absorbed in the depletion region, liberating electrons. The electric field from the reverse bias on the diode sweeps the liberated electrons across the junction, providing a current proportional to the quantity of photons absorbed. The large depletion region of the PIN diode improves the probability of absorbing photons. Intrinsic region thicknesses of approximately 4–15 μm are common. As alternatives to PIN photodiodes, avalanche photodiodes and phototransistors can also be used for the optical to electrical signal conversion. The mechanism of these devices is somewhat different from that of the PIN diode, although the output remains a current proportional to the light intensity. The main difference is that these two devices also amplify the current. Optical conversion devices have specifications that describe their performance, limitations, and parasitics. Conversion efficiency is the current produced for a given amount of light. A conversion efficiency of 0.5 A/W means that a PIN diode would produce 0.5 A of current when receiving 1 W of light. Generally, light received is in the mW or μW range, so the output current is in the mA or μA range.

Large photodiodes facilitate alignment to the fiber but have more parasitic capacitance. This parasitic capacitance can limit the frequency because it must be charged and discharged by the current produced by the photodiode. As with most high-frequency devices, capacitances should be minimized. The output from the PIN diode, current proportional to the intensity of light shining on it, is coupled into a TZA. The purpose of the TZA is to convert the current to a voltage. Con-

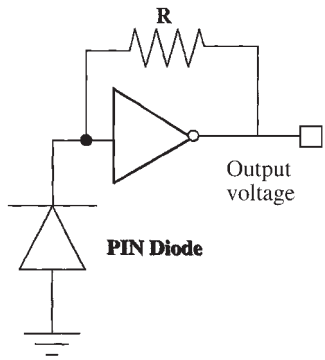


Figure 7.4 Diagram of PIN diode and TZA.

ceptually, the TZA is an operational amplifier (op amp) with a negative feedback resistor. In the ideal case, the transimpedance is the value of the feedback resistor (measured in ohms). If the PIN diode in Fig. 7.4 has a conversion efficiency of 0.5 A/W, and the incoming light is 1 mW, it will produce 500 μA of current. The op amp with a 1-kohm transimpedance resistance will convert the 500 μA to produce a voltage of 500 mV. Similarly, if the transimpedance resistance were increased to 10 kohm, the output would be increased to 5 V.

The two most important specifications for a TZA are the bandwidth and input noise level, which involve a tradeoff: Increasing bandwidth will increase noise. The key is to first choose the largest transimpedance possible to achieve the bandwidth required because the larger the transimpedance, the lower the input noise of the receiver. Of course, the larger transimpedance increases the input impedance, which reduces bandwidth. The amplifier design should use minimal current because diode shot noise increases with current. Of course, the bandwidth of the amplifier is lowered as the current is reduced. The range from the smallest detectable signal to the largest allowable signal is called the dynamic range. The largest allowable signal is the maximum that can be received without saturating the circuitry. The smallest detectable signal is determined by the input noise of the TZA. For example, if the input noise of the receiver integrated over the bandwidth is 250 nA the smallest detectable signal might be 10 times that current, or 2.5 μA. If the photodetector conversion efficiency is 0.5 A/W, the minimum difference between a low level and a high level would be 2.5 @ (*OS* A/W) = *5* μW. If this same receiver had a dynamic range spec of 30 dB, the maximum input would be *5* pW(1000) = *5* mW. A 5-mW input corresponds to an input current of 2.5 mA. If the part operates at *5* V, the maximum transimpedence gain is 2 kohm (2.5 mA × 2 kohm = *5* V). Often, a larger transimpedence is desired to provide more gain at low-input currents and to reduce noise. To achieve this, some method of gain compression can be used at high-input currents. Power supply noise rejection and substrate noise feedback are also extremely important for amplifiers with high gain and high bandwidth. If these are not properly controlled, the part will likely oscillate at low-input signal levels—perhaps being renamed the "transimpedance oscillator." Figure 7.5 is a diagram of a simple transimpedance amplifier and PIN diode. When light input to the PIN diode increases, base drive is conducted away from Q1. This starts to shut off Q1, reducing current *I1* so that the voltage across *R1* decreases, which causes the voltage on the base of QZ to rise.

Transistor *Q2* is an emitter follower that increases current drive capability. When the base of *Q2* rises, the output voltage also rises. When the output voltage rises, the current through *RZ* flows into the PIN diode. The voltage change on the output will be approximately equal to the current change in the PIN diode multiplied by the resistance of *R2*. This is the voltage change required across *R2* to replace the current being conducted by the PIN diode. Note in this diagram that

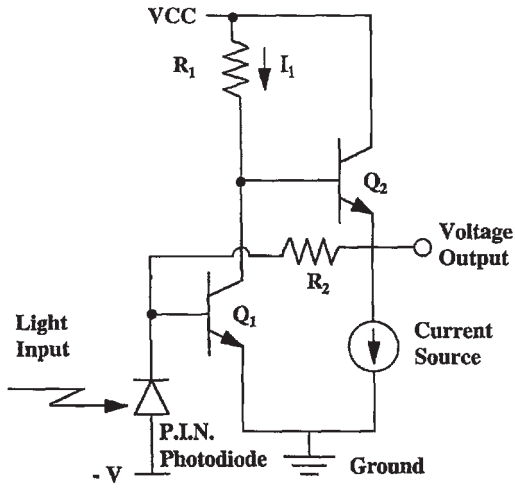


Figure 7.5 Diagram of TZA and PIN diode.

the PIN diode is biased to a negative voltage. This is necessary in this configuration because the base of Q1 will be biased to approximately 0.8 V, and the PIN diode generally requires greater than 3 V of reverse bias to operate properly. If the signal is transmitted over long distances, the optical signal is attenuated and the receiver may receive an extremely small signal. The attenuation is due to fiber loss and coupling losses at the interface between the optical components and fiber. The output voltage from the TZA must be further amplified and compared to a threshold voltage to determine if it is high or low. This is a difficult challenge, particularly if the signal has a substantial difference in the quantity of zeros and ones.

If the data are encoded for an average 50% duty cycle, an analog feedback loop similar to that shown in Fig. 7.6 can be used to set the threshold. If the input data are known to have a 50% duty cycle, output and output (OUT and OUTB) should have an average voltage in the center of the output swing. The resistors and capacitors average the output swings. If the threshold voltage is too high, OUT will tend to be low more than 50% of the time, whereas OUTB will tend to be high more than 50% of the time. The operational amplifier will sense that the noninverting input is lower than the inverting input, and the output will decrease. This will lower the threshold voltage for comparison to the TZA output by the postamplifier. This negative feedback loop will correct the threshold voltage until the outputs have a 50% duty cycle.

If the input data do not have a 50% duty cycle, the feedback circuit will still behave as if it is driving the outputs to a 50% duty cycle, and the threshold will

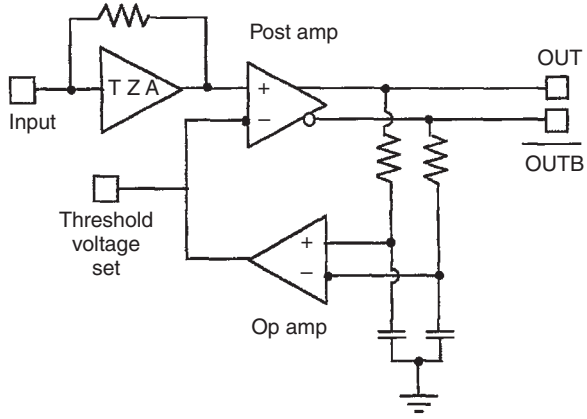


Figure 7.6 Analog feedback loop.

be incorrectly set. Setting the comparison threshold is very challenging if the data are not encoded to maintain a 50% duty cycle. A low-pass filter is often placed between the TZA and postamp. For a 1-GHz bandwidth, the TZA needs a nominal bandwidth of approximately 1.2 GHz to allow for a 20% process variation. A lot with processing such that the devices are unusually fast may have a bandwidth of 1.4 GHz, and another lot with slower devices may just meet the 1-GHz requirement. The "fast" lot will have higher noise when integrated over the 1.4-GHz bandwidth. A filter can be constructed with tight-tolerance discrete components to cut off the bandwidth at 1 GHz, eliminating the extra noise contribution from the excess bandwidth. Once the threshold is correctly set, the differential data on the outputs of the post amp are the digital equivalent of the data at the input of the laser driver. If this is one channel of a parallel system (as in Fig. 7.1a), the data are ready to use. If it is the output of a serial link (Fig. 7.1b), a clock signal has to be recovered for clocking in the incoming data, and the data need to be demultiplexed (converted from serial to parallel) to recover the parallel data that were present at the transmitter. Because all the data transitions occur at regular intervals, a phase-locked loop (PLL) can be locked to the edges for clock recovery. Once the PLL is locked to the incoming data, a clock is available that is synchronous with the data so that it can be used to clock the datastream into a register. If those data are shifted serially in a shift register and read out at the regular intervals of the clock (every eight clock cycles for an 8-bit system), the data can be converted back to a parallel datastream.

7.4. LINEAR AND LIMITING RECEIVERS

The choice of receiver design has many practical implications on the design of optical communication links. We will illustrate with an example taken from a

recent discussion within the Fibre Channel Standard membership concerning the design of receivers for 8 Gbit/s links. Two fundamental designs were considered, namely, a linear receiver and a limiting receiver.

The linear receiver consists of a photodiode (which performs optical to electrical conversion and is assumed linear with respect to incident optical power) followed by an amplifier, equalizer, and filter (which process the received electronic signal to compensate for distortion). Instead of using a simple comparator in limiting receivers, the output of the receiver remains linear or proportional to the optical signal. The limiting receiver output is either high or low and "limited" by the voltage output of the comparator.

Since the linear receiver can be designed to perform electronic dispersion compensation (EDC), it is possible to achieve better signal integrity over longer transmission distances using this approach. Typically, EDC can roughly double the achievable link distance. This is an important advantage, since link distance is typically reduced by half for every doubling in data rate. Without this approach, 8 Gbit/s links would have been limited to under 100 meters and might have been unable to meet customer distance requirements. Unfortunately, standardizing EDC is a significant new venture that adds complexity to the receiver design and possibly cost due to low volume. At the time 8 Gbit/s links were being developed, it was felt that the datacom market would adopt the new 8 Gbit/s interface more favorably if the hardware cost was near parity with a 4-Gbit/s link (doubling the bandwidth at very low incremental cost). To achieve this, many designers opted for the alternative and somewhat simpler limiting receiver design.

As illustrated schematically in Fig. 7.7, the limiting amplifier stage is used to boost the variable amplitude of the preceding amplification stage to a constant, "limited" amplitude for subsequent clock and data recovery and decision thresholding (the actual circuit may consist of several stages). This design is nonlinear due to its high limiting gain. Note that the limiting amplifier has a low-pass transfer function; thus the receiver circuit sensitivity is determined by the noise

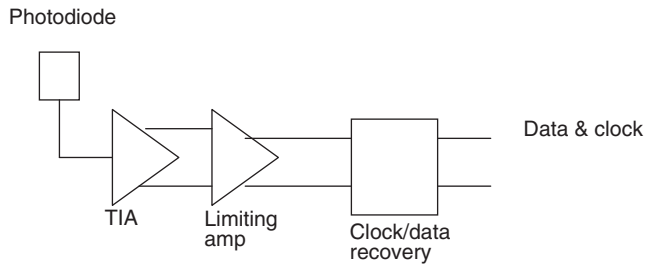


Figure 7.7 Schematic of limiting receiver design.

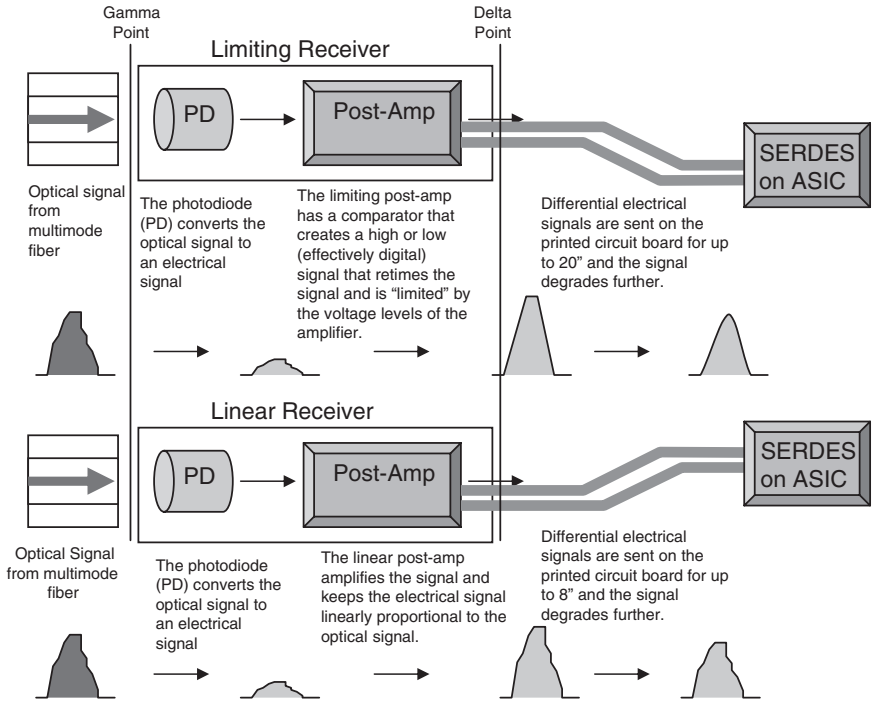

Gamma Point
Delta Point
Limiting Receiver
PD
Post-Amp
SERDES on ASIC
Optical signal from multimode fiber
The photodiode (PD) converts the optical signal to an electrical signal
The limiting post-amp has a comparator that creates a high or low (effectively digital) signal that retimes the signal and is "limited" by the voltage levels of the amplifier.
Differential electrical signals are sent on the printed circuit board for up to 20" and the signal degrades further.
Linear Receiver
PD
Post-Amp
SERDES on ASIC
Optical Signal from multimode fiber
The photodiode (PD) converts the optical signal to an electrical signal
The linear post-amp amplifies the signal and keeps the electrical signal linearly proportional to the optical signal.
Differential electrical signals are sent on the printed circuit board for up to 8" and the signal degrades further.

Figure 7.8 Comparison between linear and limiting receiver designs; gamma and delta points represent measurement locations in the receiver signal path.

integrated within the receiver input bandwidth. Narrowing the bandwidth in an effort to reduce noise also reduces the effective receiver sensitivity. Furthermore, the limiting amp bandwidth determines the rise and fall times. Generally, shorter rise/fall times (required for high data rates) imply a larger receiver output bandwidth. There is thus an inherent tradeoff between the speed and sensitivity of the limiting amp design.

Although some design approaches can achieve reasonably good combinations of sensitivity and speed, the lack of signal equalization means that limiting receivers must be used on shorter link distances. The resulting 8-Gbit/s Fibre Channel standard for short-wave transmitters on legacy (OM2) fiber is thus limited to a maximum of 50 meters, whereas more expensive, complex receivers could have achieved 150 meters or more. Applications that require longer distance must either re-cable their infrastructure with higher bandwidth fiber or reconfigure their network to reduce the maximum unrepeated distance to 50 meters or less. Recently, it has been reported [1] that limiting amplifiers can be achieved with

11-GHz input bandwidth, 14-ps rise/fall times, 14-mV sensitivity, and about 39-dB limited gain, at the cost of higher DC power consumption (430 mW).

As shown in Fig. 7.8, the main interfaces to the limiting and linear transceivers are defined at the gamma and delta points. The gamma point is the optical signal requirements at the input to the transceiver. FC-PI-4 [2] defines the gamma points for both linear and limiting gamma points, and they are slightly different. The main differences between linear and limiting transceivers are at the electrical outputs from the receiver, known as the delta points. The delta points define the electrical output signal properties. The limiting delta point output is basically digital, while the linear signal is analog. The gamma and delta points of 8 Gb/s limiting transceivers are defined by FC-PI-4, while the linear points are defined by FC-PI-4 and SFF-8431 (SFP+) [3].

The linear solutions use EDC technologies to extend the length of the link. A multitap filter is used to compensate for signal degradation. SFF-8431 defines the electrical signal characteristics for 10 Gigabit Ethernet in addition to 8G Fibre Channel. In addition, the extra link budget that EDC enables may relax requirements on the transmitter and receiver. This can be significant, since the transmitter has traditionally been the highest cost component in the transceiver.

REFERENCES

1. Titus, W. 2007. Proc. 2007 IEEE Sarnoff Symposium, Princeton, N.J.
2. http://www.t11.org/t11/docreg.nsf/ufile/07-255v1, Fibre Channel—Physical Interfaces -2 (FC-PI-4), Hossein Hashemi.
3. ftp://ftp.seagate.com/sff/SFF-8431.pDF, SFF-8431 Specification for Enhanced 8.5 and 10 Gigabit Small Form Factor Pluggable Module “SFP+”, Ali Ghiasi.

This page intentionally left blank

8

Optical Subassemblies

Herwig Stange
Infineon Technologies, Fiber Optics, Berlin, Germany

8.1. FUNCTION OF THE OPTICAL SUBASSEMBLY

Within the fiber-optic link, the optical subassembly converts the data signal from an electrical current to optical radiation in the fiber and vice versa. The optical subassembly comprises the electro–optical converter, optic, and connector components (Fig. 8.1).

Key components of the optical subassemblies are the electro-optic active elements. In the case of the transmitter, this is the light-emitting diode (LED) or the laser, edge-emitting laser diode (EELD), or vertical cavity surface-emitting laser (VCSEL). The optical subassembly of the receiver comprises the photodetector (PD), mostly a positive-intrinsic-negative (PIN) diode or an avalanche photodiode (APD). These elements are the actual electro-optical converters within the optical subassembly [1].

The optics of the optical subassembly couples the radiation from the electro-optic active element into the fiber. In many optical subassemblies the optics is a lens element. At the transmitter side this element collects the emitted radiation of the laser or LED and focuses it into the fiber. At the other side of the fiber-optic link at the receiver, the optics couples the divergent output beam of the fiber into the photodiode.

Finally, the fiber has to be fixed to the optical subassembly. Optical subassemblies with a pigtail have a permanently fixed fiber mostly with a connector at the end. Other optical subassemblies have a receptacle. This is a mechanical socket for a connector. The receptacle guides the connector with the fiber to the coupling position.

There are different names for the optical subassembly. Sometimes it is called a coupling unit, emphasizing the function of optical coupling. But mostly the name “optical subassembly” or the short form “OSA” is used.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2002, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

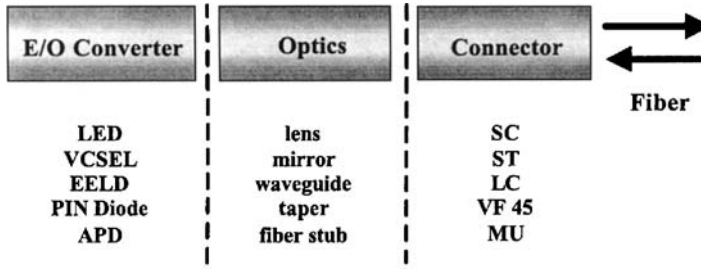


Figure 8.1 Structure of an optical subassembly.

8.2. BASIC PROPERTIES OF THE TRANSMITTER OSA

The main function of the transmitter OSA is to convert current to radiation, which is launched into the optical fiber. The key parameters describing this property of the transmitter OSA are the efficiency of the electro-optic conversion and the optical coupling. This is the slope efficiency or the differential quantum efficiency of the laser combined with the losses of the optical coupling. The dimension of the efficiency is watt per ampere (W/A). That implies that the differential quantum efficiency of the laser or LED chip defines the upper limit for the total efficiency of the transmitter OSA. In particular, the optical coupling efficiency for LEDs can be very low due to the highly divergent radiation emitted by the LED. Also, the optical coupling of edge-emitting lasers with single-mode (SM) fibers can result in strong losses. This type of laser with an emission wavelength of 1300 nm or 1550 nm is used for intermediate and long-reach datacom applications that demand more launched power than short-haul applications.

The electro-optic active element mainly defines the parameters for the specification of the OSA. The wavelength of the emitted radiation of the OSAs is given by the spectral output of the semiconductor device. Antireflection coatings at the optical surfaces can reduce the loss due to reflection and optimize the coupling efficiency of the OSA. A further reason to reduce reflections back to the laser is that they can disturb the laser emission and increase the output noise. Another possibility to prevent backreflections is to avoid optical surfaces exactly perpendicular to the laser beam. The relative intensity noise (RIN) in db/Hz measured to the average optical power characterizes the noise of the laser, which should be measured with a defined return loss. This parameter represents the reflections from connectors and the receiver. For example, the Fibre Channel Standard [2] demands a measurement at 12-dB return loss.

The optical rise and fall time of the transmitter OSA is mainly determined by the properties of the laser or LED chips. But again the integration of the chip within the OSA package can detrimentally affect these parameters. Parasitic

capacitance and inductance due to long lines can affect the speed of the device electrically.

Threshold and differential quantum efficiency of lasers change with temperature and age. However, the fiber link demands a constant power. The problem is mostly solved by controlling the laser diode. A photodiode within the transmitter OSA detects a part of the laser radiation. This monitor diode current has to be proportional to the power launched into the fiber in order to control this value. Therefore, a constant ratio between the optical coupling of the fiber and the optical coupling of the monitor diode is necessary. The launched optical power should remain constant over the operating temperature range by controlling the monitor current. The tracking error describes the maximum deviation over temperature and is the ratio of minimum or maximum power (based on the maximum difference) and the power at the reference temperature in units of dB. Mostly it refers to the power at 25°C room temperature.

8.3. BASIC PROPERTIES OF THE RECEIVER OSA

The responsivity of the receiver OSA summarizes the efficiency of the electro-optic conversion of the detector and the efficiency of the optical coupling between fiber and detector. It is the ratio of generated current to optical power with the dimension ampere per watt (A/W). Due to the spectral sensitivity of the detector, it depends on the wavelength of the incident radiation. Many OSAs have an anti-reflection coating at the optical surfaces in order to improve the optical coupling and therefore the responsivity.

Not only a high transmission but a low reflectance of the receiver OSA is important. Reflected light could disturb the laser and can cause an increase of laser noise. A high return loss of the receiver prevents strong backreflection to the laser of the transmitter. The return loss is defined as the ratio of incident optical power to the reflected optical power in units of dB.

Often the receiver OSA includes a preamplifier. The advantage of integrating the preamplifier in the OSA package is that this design enables a short connection between detector and amplifier. This is important because the low currents of the detector are very sensitive to electrical crosstalk. A package like a transistor outline (TO) can efficiently protect this connection. In addition, parasitic capacitance and inductance due to lead-throughs of the package would affect the photodetector and reduce the bandwidth of the receiver.

8.4. COUPLING RADIATION FROM A LASER DIODE INTO A FIBER

For most applications, the first design goal is to launch the radiation from the laser or LED into the fiber as efficiently as possible with least cost. The optics of the OSA has to match the diameter and beam divergence of the emitter with the

fiber core diameter and the numerical aperture of the fiber. An area mismatch occurs if the beam diameter is larger than the core of the fiber. Reducing the distance between the emitter and the fiber reduces the loss due to the area mismatch. Positioning the fiber directly in front of the emitter surface (butt coupling) reduces the effect of the divergence of the beam to a minimum. Even if the beam diameter is smaller than the fiber core, losses can occur due to numerical aperture mismatch. A part of the beam exceeds the acceptance angle of the fiber and is not guided by the fiber [3,4].

Besides the optical coupling efficiency, other requirements may have an impact on the optical design. The OSA or at least the transmitter has to be eye safe. A laser product classification is preferred that imposes the fewest constraints on the user of the transmitter (see Chapter 6). One approach is to have divergent emission, which reduces the energy density of the accessible radiation reaching the eye from the open transmitter port [5].

Some transmitters are used alternatively for single-mode fibers and for multimode graded-index fibers. It is possible to increase the bandwidth of the multimode fiber link by reducing the differential mode delay (DMD) in the multimode fiber. The optical design of the OSA is one approach to minimize the differential mode delay; another is the design of the multimode fiber profile itself. The optical coupling of laser and multimode fiber has a strong impact on the distribution of modes. The optical design can optimize the generation of an appropriate mode distribution in the multimode fiber.

Coupling an edge-emitting laser with a beam divergence of 30° into a single-mode fiber with an accepting angle smaller than 6° demands an optic that transforms beam diameter and divergence. A magnifying lens shapes the beam into a slightly convergent beam that matches with the numerical aperture of the single-mode fiber.

One advantage of vertical cavity surface-emitting lasers is that fiber coupling is simplified. VCSELs are usually designed with a circular aperture, resulting in a circular beam. This, combined with the low divergence of the emitted radiation, implies an improved fiber launch efficiency compared to the edge-emitting components. Figure 8.2 shows the ray trace modeling of a VCSEL coupling with an aspherical plastic lens.

This lens can be integrated in a plastic flange. The plastic flange has to fulfill very different requirements. It has to be optical transparent and has to have surfaces in good optical quality at the lens. The tolerances of the inner diameter of the flange have to be very low. These properties should be stable at high temperatures. To date, the precision achieved is only sufficient for multimode applications.

Figure 8.3 shows the TO can and plastic flange fixed together by epoxies. The VCSEL in the TO can is adjusted to the optimal position before the epoxy is cured. In order to reduce production time light-curing epoxies are used.

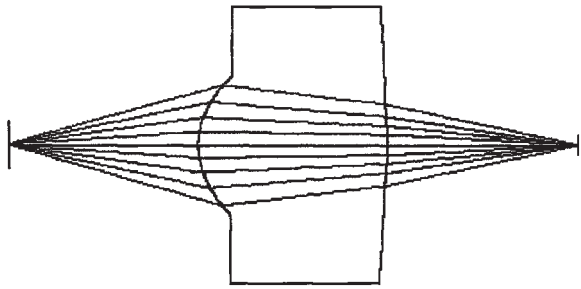


Figure 8.2 Ray trace modeling of VCSEL coupling.

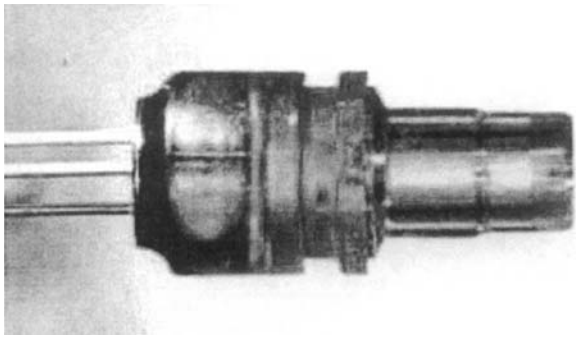


Figure 8.3 VCSEL-OSA with plastic flange.

A crucial parameter of all receptacles is the inner diameter of the flange. The tolerance of the connector in the flange causes a variation of launched power. Figure 8.4 shows two different cases of coupling into a single-mode fiber. The variation of coupled power is very low if the receptacle is exactly aligned at the maximum of the coupled power. A tolerance of +/–1 µm of the connector position causes a power variation of only a few percent. The same tolerance of +/–1 µm leads to strong variation in launched power if the alignment is not perfect and the flange is de-adjusted. Figure 8.4 shows a de-adjustment of 4 µm. In this case, the single-plug repeatability is poor. It is defined by the variation of coupled power after several matings with the same connector.

The cause of de-adjustment is not only misalignment but also distortion during gluing or thermal expansion during operation. These effects can be minimized by designing a highly symmetrical component.

For applications with single-mode fibers, materials such as ceramics and stainless steel are mostly preferred. These materials can be manufactured with high

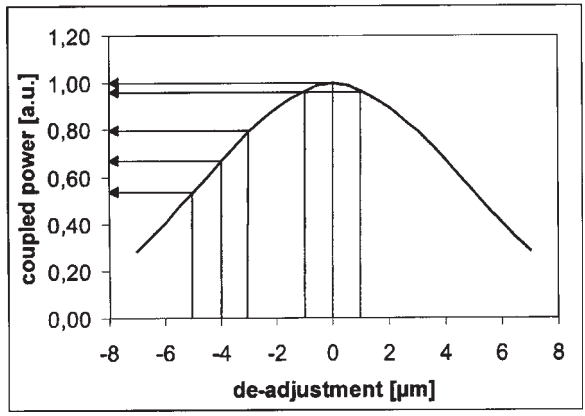


Figure 8.4 Coupling curve.

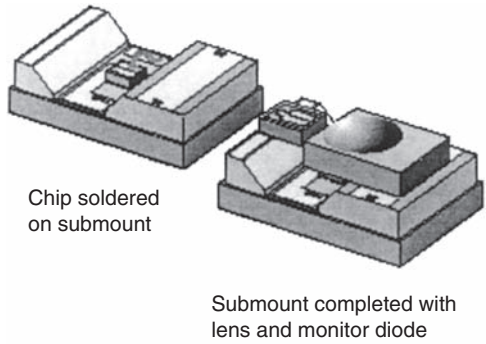


Figure 8.5 Laser submount module.

precision and have low coefficients of expansion. Steel parts can be welded almost free of distortion.

The tolerance of the connector position is given by the difference of inner flange diameter and outer connector diameter and the different eccentricities of connectors. These tolerances have to be considered if the cross plug repeatability shall be determined. The cross plug repeatability is a property of both the OSA and the connectors.

Edge-emitting lasers for 1300-nm or 1500-nm application demand a different assembly. Figure 8.5 shows a compact assembly on a Si submount including a lens and a monitor diode. The laser chip is placed at the center of a Si submount with two glass prisms on either side that reflect the front and rear emitted beams of the laser to the top. The laser beam emitted at the front facet of the laser chip

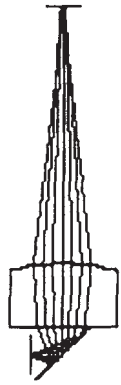


Figure 8.6 Beam path of the laser submount module.

is focused by a Si lens that is on top of the right glass prism (Fig. 8.6). On the left side, the rear beam is reflected to the monitor diode mounted on the left glass prism.

The hybrid integrated fiber-optic module on a silicon submount chip allows a very compact design that brings the focusing lens almost directly in front of the laser chip. The total length of the beam path can be very short. The distance between laser chip and fiber or connector can be kept short, minimizing the size of the OSA.

Using standard semiconductor wafer technology allows efficient volume production of laser modules. Anodic bond processes and solder bonding techniques guarantee stable assemblies with high reliability. The assembly of the complete active fiber-optic component on a wafer has further advantages. Burn-in and testing of the laser module can be done with the fully assembled module on the submount wafer. Figure 8.7 shows the wafer before the last production process, the assembly of the monitor diode [6].

8.5. COUPLING RADIATION FROM A FIBER INTO A PHOTODETECTOR

On the other side of the optical link the light emitted out of the fiber has to be focused on the photodetector. The beam divergence is determined by the numerical aperture of the fiber. A 62.5-μm graded-index fiber with a numerical aperture of 0.275 results in a half angle of 16°. At this angle the power intensity has decreased to 5% of the maximum power intensity at the center of the far field. This diverging beam has to be focused in the OSA on the photodetector. Many applications allow a sensitive area that is bigger than the core diameter of the fiber.

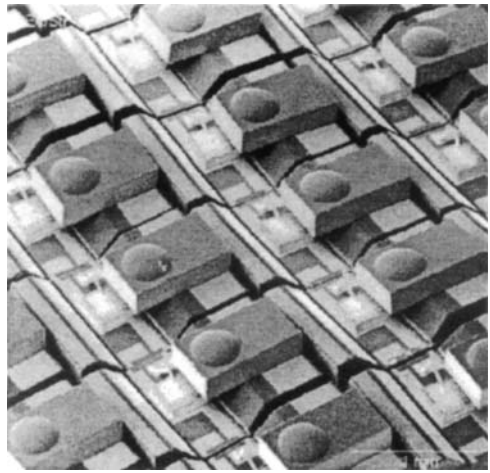


Figure 8.7 Bonded laser diodes on Si submount wafer with adjusted Si lens.

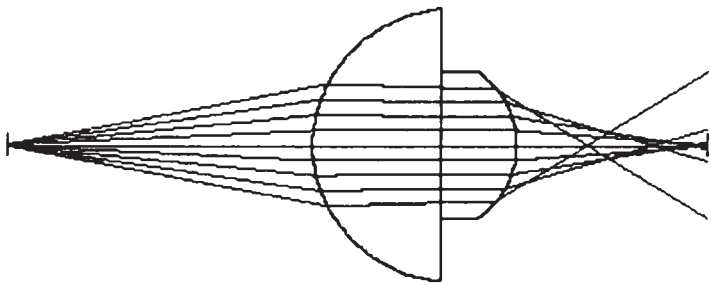


Figure 8.8 Ray trace model of a TO lens focusing the light on the detector area.

Therefore most optics achieve a sufficient coupling of light on the photodetector. The sensitive area of the photodetector is forced to become small only for high data rate applications in order to move the cutoff frequency of the device to high frequencies.

Contrary to the transmitter OSA, the optical coupling with a single-mode fiber is easier due to the lower numerical aperture of 0.1 and the smaller core diameter of about 9 µm. The half angle of the far field is below 6°.

Figure 8.8 shows the lens of a TO cap focusing the radiation out of the fiber onto the detector. These lenses can be produced in a simple melting process. The focal length varies strongly. Nevertheless, for many applications this component is sufficient. With the same assembly design as described for the laser, the

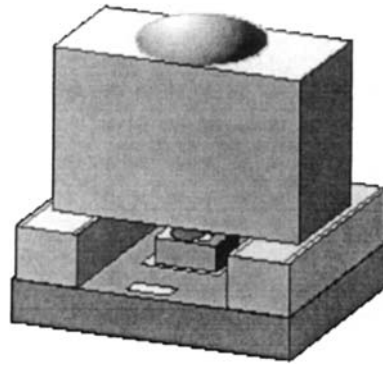


Figure 8.9 Receiver micromodule.

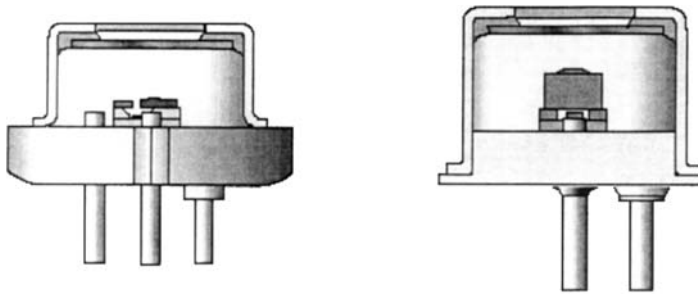


Figure 8.10 TO can housing of a laser and a detector micromodule.

assembly of a very compact design is possible. The small distance between lens and detector allows a short length of the total beam path. The PIN diode is at the center of the submount between two posts that carry the Si lens (Fig. 8.9).

8.6. PACKAGING OF OPTICAL SUBASSEMBLIES

The TO can is a standard housing with different functions. It protects the sensitive chips from mechanical damage. The hermiticity of the housing is good protection against environmental impacts such as humidity or an aggressive atmosphere. In particular, ionic impurities in a humid atmosphere can cause severe damage to the chips. The header of the TO can dissipates the heat of the device arising from the laser and reduces the thermal heating (Fig. 8.10).

The beam is emitted through a glass window. It can be coated with an antireflection coating to optimize the optical coupling and to prevent disturbing reflections back to the laser.

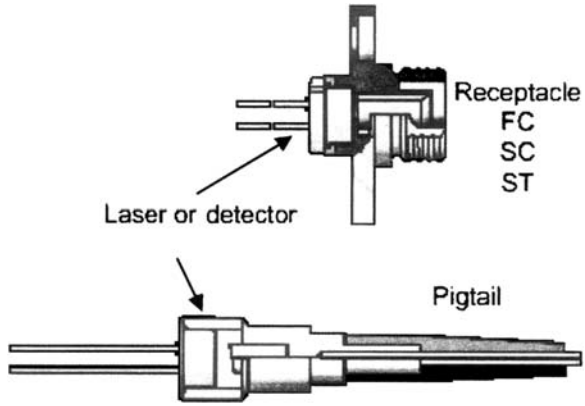


Figure 8.11 Optical subassemblies with TO cans.

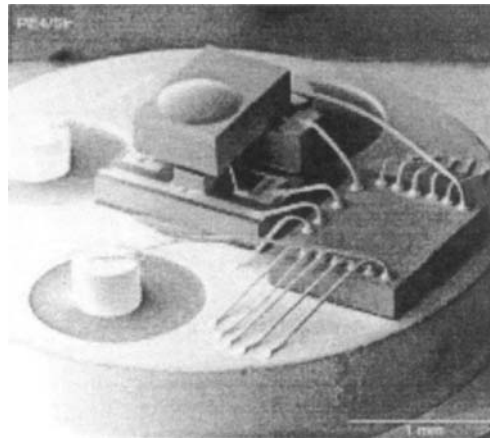


Figure 8.12 PIN diode with Si lens and preamplifier on TO header.

The metal parts of the TO can allow highly stable welding processes. The tolerances for adjustment are in the range of 1 µm. Figure 8.11 shows different assemblies. The receptacle OSA combines a TO can with a flange that fits with a fiber connector. The second is a pigtail OSA with a connectorized fiber. Typically, the first one is used for short-range datacom applications that demand pluggable connectors at the device. The pigtail solution is favored for long-range telecom applications that demand high optical coupling efficiencies with low return losses. This is achieved by the slightly tilted surface of the pigtail fiber in front of the laser element.

The TO package allows us to integrate the preamplifier into the TO can (see Fig. 8.12). This saves space within the transceiver because otherwise the pream-

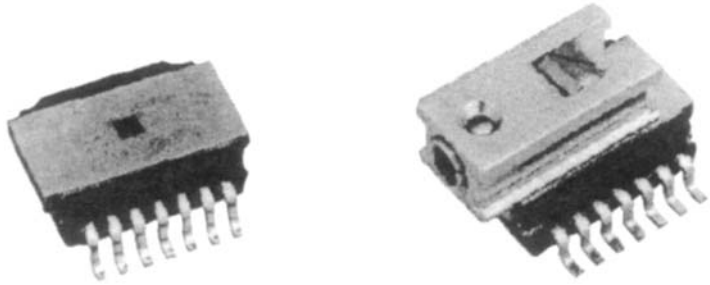


Figure 8.13 Optical subassemblies in SMD package.

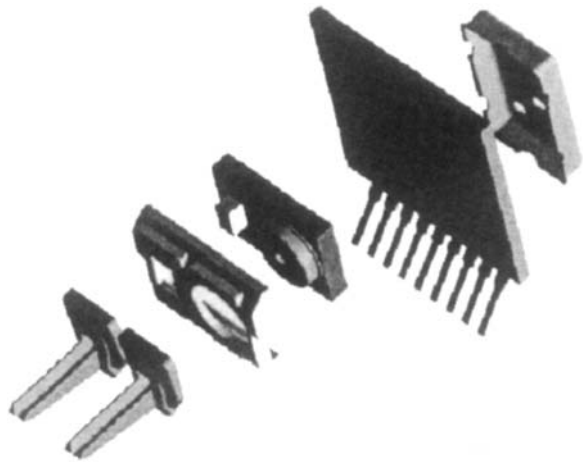


Figure 8.14 PCB assembly.

plifier had to be placed on the printed circuit board of the transceiver. In addition, the TO can builds an efficient shielding against electromagnetic emission and makes a short connection between photodetector and preamplifier possible (see Section 8.3).

Instead of packaging the electrooptic components in a TO can, a micromodule can be assembled on a lead frame (Fig. 8.13). A transfer molding process encapsulates the surface mount device. The SMD technology allows large-scale fabrication at low cost.

Figure 8.14 shows a different packaging. The micromodules (center part) with the electro-optic components are assembled onto the printed circuit board (PCB). Small metal caps prevent distortions from electromagnetic emission. The part shown in Fig. 8.14 was built for a VF45-connector. This connector guides the fibers of the connector via v-grooves to the coupling position. The optical

adjustment of the v-grooves is fixed by light-curing epoxies. Adjustment tolerances of 2 µm can be achieved. Symmetrical design is important in order to prevent changes due to the shrinking epoxy. The advantage of this process is that it can be done with simple equipment that does not demand a high investment.

8.7. OPTICAL SUBASSEMBLIES FOR PARALLEL OPTICAL LINKS

One approach to reduce the size of fiber-optic link modules is to bundle optical channels. Parallel optical links allow the miniaturization on the board and allow cost reduction (see Chapter 11). Figure 8.15 shows an example of how the number of channels can be increased on a single chip [7].

The edge-emitting laser array is mounted on a submount. High-accuracy Si micro-bench technology achieves the precise spacing necessary to integrate a number of optical links at this level.

VCSEL arrays can be coupled with the fiber bundle without focusing optics and where the fiber axis is parallel to the PCB. The laser radiation emitted from the top surface is reflected into the axis of the fibers by the 45° angled surface (Fig. 8.16).

The top view of the device (Fig. 8.17) shows the VCSEL array in the center and the plastic part with zigzag profile holding the fiber array. This part is made of a highly filled plastic material that is very temperature stable.

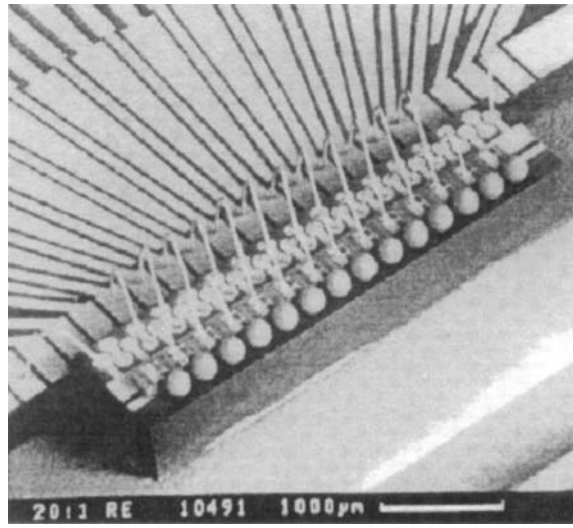


Figure 8.15 Laser array with lenses.

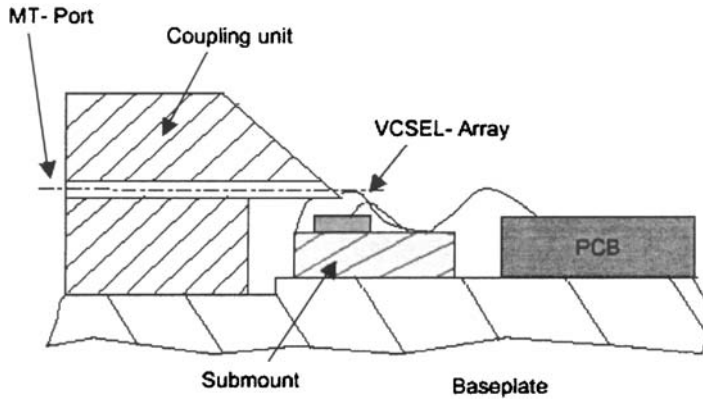


Figure 8.16 Coupling design of a parallel optical link with VCSEL array.

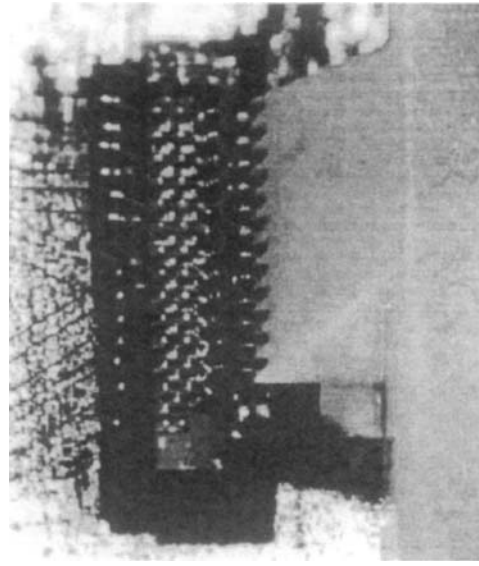


Figure 8.17 Top view at VCSEL array (center) and high-precision plastic part with fiber array (right).

After assembly, a nonhermetic sealing is applied for protection against mechanical forces and environmental impacts.

8.8. OUTLOOK

The goal of future developments is clear and can be easily extrapolated from the past. Fiber-optic links will need optical subassemblies that enable transmission at increasingly higher data rates. The costs of the devices have to be reduced, the size will shrink, and the power consumption will decrease.

There is a strong demand to transmit more and more data with smaller and smaller devices. In particular, the cross-sectional dimension of the optical port has to be minimized, because the accessible front or rear panels of devices such as routers is a limiting factor. This leads directly to the requirement of small cross-sectional dimensions of optical subassemblies with electrical connections that are able to transfer the demanded high-frequency signal. The shrinkage of the transceiver port size for a duplex-connector is indicated in Fig. 8.18 by the size of the black rectangles. Within the last ten years the port area could be reduced by a factor of 5 whereas the data rate increased by a factor of 10, resulting in the strong increase of the data transfer rate per area.

There will be different solutions for different applications from short-reach to long-haul data communication. The merging of data and telecommunication will

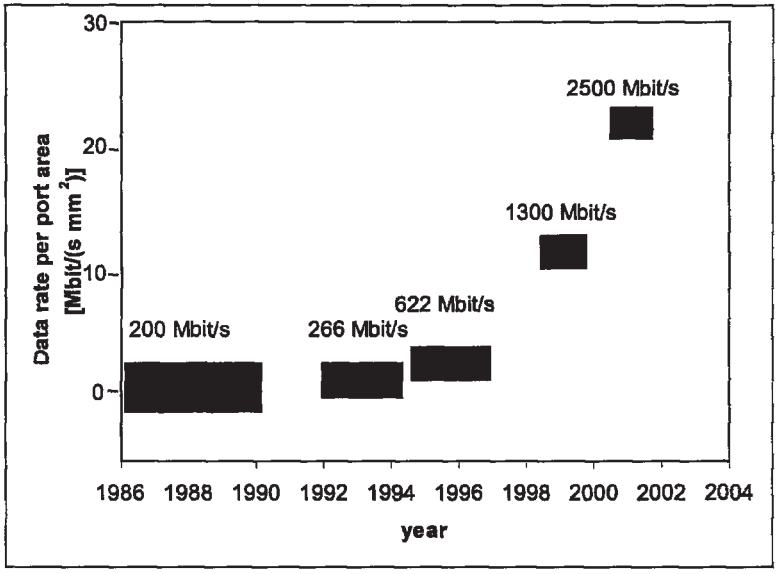


Figure 8.18 Development of data rate and port area.

continue. Technologies from telecommunication such as wavelength-division multiplexing (WDM) or bidirectional transmission on a single fiber will also be applied in datacom applications.

There are a variety of packaging technologies for optical subassemblies. For many applications OSAs with TO canned diodes and SMD-OSAs are beneficial as well as microbench technology with a fiber coupling between diode and fiber on submount level. In the future the monolithic integration on the IC gives the possibility of further miniaturization. Today all these technologies are used. Depending on applications with different requirements and the volume of production, each technology has and will have an importance.

ACKNOWLEDGMENTS

Thanks to my colleagues at Infineon Technologies for supporting me in so many ways. In particular, thanks to Michael Langenwalter for coordinating our contributions. Special thanks to Thomas Murphy for providing me with many helpful suggestions for improvement. Thanks also to Klaus Schulz for fruitful discussions and to Rainer Ulrich for the many photographs he provided.

REFERENCES

1. Agraval, Govind. 1997. *Fiber-optic communication systems*, 2nd ed. (Wiley series in microwave and optical engineering.) New York: John Wiley & Sons.
2. ANSI X3T9.x and T11.x, Fibre Channel (FC) Standards incl. FDDI, SBCON and HIPPI-6400, at http://web.ansi.org/default.htm.
3. Born, M., and E. Wolf. 1986. *Principles of optics*, 7th ed. Cambridge: Cambridge University Press.
4. Pedrotti, F. and L. Pedrotti. 1993. *Introduction to optics*, 2nd ed. Upper Saddle River, N.J.: Prentice Hall.
5. International Standard IEC 60825-1, 1993 incl. Amendment 2, January 2001, ISBN 2-8318-5589-6, Safety of laser products—Part 1: Equipment classification, requirements and user's guide.
6. Althaus, H., W. Gramann, and K. Panzer. 1998. Microsystems and wafer processes for volume production of highly reliable fiber optic components for telecom- and datacom-application. *IEEE Transactions on Components, Packaging, and Manufacturing Technology—Part B*. 21:2.
7. Karstensen, H., et al. 2000. Module packaging for high-speed serial and parallel transmission. ECTC 2000, Las Vegas, Nev.

This page intentionally left blank

9

Alignment Metrology and Manufacturing

Darrin P. Clement
Maponics, East Thetford, Vermont

Ronald C. Lasky
Consultant, Medway, Massachusetts

Daniel Baldwin
Manufacturing Research Center; School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia

9.1. INTRODUCTION

This chapter provides some benchmarks to use when determining characteristics of coupling in actual components. The theory behind such data will only be presented in the references. Alignment will be discussed in both optical and mechanical contexts. We will assume that the reader has a familiarity with the concepts herein and that this chapter will be used primarily as a reference.

We will then discuss the current techniques for achieving reliable mating between optical components for data communication. Active manual, active automated, and passive automated processes will be explored with an emphasis on expected performance and relative costs. Throughout the chapter, we will not discuss telecommunications applications except as they may apply to data communication in the present or near future.

9.2. INTERFACE DEFINITION AND IMPORTANCE

An optical "interface" includes several interconnection elements. The cases considered here involve the mating of an optical source to an optical fiber and/or the coupling between a fiber and a photodetector. A "source" may be either a laser diode (LD), a light-emitting diode (LED), or a vertical cavity surface-emitting laser (VCSEL). A lens is typically situated between the source and the fiber waveguide to increase coupling efficiency.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

A packaging structure for datacom contains all of these elements and provides the required degree of precise optomechanical alignment among all elements. This structure is an optical subassembly (OSA), and we distinguish those for transmitter and receiver ends of the link as TOSAs and ROSAs, respectively. With the optoelectronic elements packaged in this way, the problem becomes an optomechanical one of aligning the fibers.

When using an LED, the fiber into which light is coupled would be multimode, whereas single-mode fibers require LD sources. In either case, the fibers used must be ferruled in order to provide durability and accuracy in optomechanical alignment. These cylindrically ferruled fibers are joined to the OSA module by "plugging" them into a bore (an alignment sleeve) located in the OSA. The ability of the ferrule to stay within the bore is characterized through spring and retention forces. Typically accepted values are 7.8–11.8 N for the ferrule spring force and 2.9–5.9 N for the bore retention force [1, 2].

A popular design scheme for datacom is the "connectorized" module. Connectorized in this context means the provision of an optomechanical construction at the module that can provide secure alignment and retention of the fiber while also allowing simple manual detachment and reconnection.

Generally, the trend in datacom hardware is to employ "duplex" connectors/modules, meaning that the connector hardware accommodates a pair of fibers, one each for the sending and receiving paths. Thus, at the transceiver module, the TOSA and ROSA are situated alongside each other, and both fibers can be connected or detached in the same manual operation. The fiber distributed data interface, Fibre Channel Standard (a derivative of SC), and the Enterprise System Connection series of duplex connectors/modules are examples that embody this design [3]. Ribbon connectors are used for array devices whereby multiple fibers (usually multimode) are coupled into a linear (one-dimensional) array of lasers or detectors. These are all discussed in Chapter 1. To date, several new connector types are under development for use with VCSEL array transmitters.

9.3. LIGHT COUPLING

9.3.1. Min/Max Coupled Power Limitations

Because we are concerned with the pluggability of fiber-optic connectors into OSAs, both mechanical and optical characteristics must be considered when determining the amount of coupling deviation acceptable. In order to achieve successful data transmission, criteria must be established that put a lower limit on what is considered "enough" coupled power. The coupled power range (CPR) is a measure of how much variation in light transfer is allowed when mating any given fiber with any given transmitter optical subassembly to ensure adequate system performance and yet not violate laser safety requirements.

Definition. Coupled power range is the allowable difference between the minimum and maximum allowed power.

In order to account for all the losses along a system's length, a concept referred to as the "link budget" must be understood and has been explained in Chapter 7.

Whenever a receiver is said to have a sensitivity value, it must also be understood that this is with respect to an attainable bit error rate (BER). Data communication often requires a BER of 10^{-15} which means that only 1 in 10^{15} bits can be erroneous. Telecommunication requirements are more relaxed and can be as low as BER = 10^{-9}. Data links have specifications midway between the two values (e.g., 10^{-12}). See Ref. [4] for a tradeoff analysis.

For example, a series of components (fiber connectors, fiber length, couplers, etc.) could result in an approximately 13-dB total link loss. The minimum light that a typical photodetector can receive for a successful BER of 10^{-15} is approximately –23 dBm. This leaves a minimum launching power of –10 dBm. Thus, the TOSA and connector must achieve at least –10 dBm of coupling every time (e.g., within a 3 σ limit, or less than 27 faulty connections per 10,000). These numbers will vary depending on particular system implementations.

Maximum coupled power is a strong concern in laser systems (compared to LED systems) simply because of the higher available intensities. Some local area networks are user accessible, meaning that nontechnicians are able to connect/disconnect optical fiber cables between computers. This access provides robustness but also presents potential safety hazards to the user. In addition, there is concern that technicians may be exposed to focused amplified light when inspecting an optical link for problems. International laser safety standards place a limit of less than 0.600 mW (–2.2 dBm) of optical power in the link [5], whereas the American National Standards Institute standard requires less than 100 s of exposure to 1 mW of light at $\lambda = 1.3$ or 1.55 μm [6]. See Chapters 7 and 8 for more information regarding planning a link and safety considerations.

This situation is unique to datacom because the telecom industry typically has user-inaccessible optical fibers. Thus, if the telecommunication power received at a detector is too weak, the laser output can simply be increased with much less concern for safety. Their limit of maximum coupled power is more likely to be dictated by the characteristics of the components. Recent developments have provided a different perspective on laser safety concerns. The alignment of source to fiber need not be controlled so as to keep the launched power below a certain value. Automatic power controls and open fiber controls, which electronically adjust the power while monitoring the transmission, are permissible under International Electrotechnical Commission (IEC) rules.

Another reason to have limits on output power, even with LED sources, is receiver saturation. Receiver sensitivity can be very high, approaching –40 dBm

or better, but the dynamic range of the receiver may be limited. If the power at the detector is too high and saturation occurs, response speed can be degraded to a point at which link performance may be compromised.

9.3.2. Multimode vs. Single-Mode Connections

We should at this point contrast the cases of single- and multimode. In doing so, we will necessarily have to compare the situation of laser sources with that of LED sources. Although a laser can be used with either fiber type, LEDs are for all practical purposes suitable only for multimode fiber. Let us assume two hardware sets, one for single mode (SM) and one for multimode (MM), that have the same absolute tolerances on dimensions, angles, and distances. For example, a loss of 1 dB in a SM system could occur with a transverse displacement of the fiber core by approximately 1 μm. It takes approximately 15 pm of similar multimode displacement (50-μm diameter core) to give rise to the same fractional loss of power [7]. The analysis of coupling geometry and efficiency is more difficult for the multimode case; however, connector hardware parts that fail optomechanical tolerances for single mode are often serviceable for multimode applications. See Refs. [8–14] for details.

Multimode fiber used in conjunction with LED sources can be very low cost. Currently, there are a number of commercial offerings of fiber-optic transceivers designed for use with multimode fiber that provide distance-bandwidth products in the hundreds of Mb/km. Requirements for precision in the critical dimensions of the multimode connector interface are readily achievable with good- to high-quality machining. In certain applications, even a plastic fiber with core diameter of 100 μm may suffice.

We must bear in mind that the discussion has emphasized the need for short-distance data transmission at high data rates. For much longer distances, single-mode fiber should be used to ensure high bit rates in excess of 1 Gb/s. For this reason, the long-term strategy may be to use single-mode fiber “to the curb” supplanted by multimode fiber “to the home or office.” Although MM fiber is no cheaper than SM fiber, it has been possible to reduce component cost with multimode solutions.

However, many companies are now marketing an “array” type of transmitted receiver set. In these parallel transmission systems, an array of VCSELs are coupled into a ribbon fiber. The highly directional VCSELs coupled into multimode fibers provide relaxed mechanical tolerances in order to achieve consistent coupling.

9.4. ELEMENTS OF COUPLED POWER

The practical coupling considerations for datacom connections are outlined in this section. Coupling to the TOSA and ROSA requires different considerations. The large active area of a typical photodetector and the larger MM fiber core

diameter result in the fiber/ROSA interface having much less stringent optomechanical constraints than the TOSA/fiber interface. Thus, the TOSA/fiber interface will be emphasized in this chapter.

Poor control of the many tolerances involved can cause two basic problems in connector/module mating. The first problem may be observed when one plugs the same connector into the same module and obtains extensive variations in the launched power. This variation is referred to as plug repeatability (PR). Two decibels is typically the most extensive acceptable PR variation.

Definition. Plug repeatability is the variation in coupled power for multiple connections between the same components.

The second variation is referred to as cross-plug range (XPR) and is measured by subtracting the lowest power reading from the highest power reading in a large sample of different connectors mated to the same module. A large cross-plug range is of concern because it is a significant source of the variation of the power launched to any given link. In many cases, improving the "peak" performance is less important than ensuring a small change in performance when interchanging connections.

Definition. Cross-plug range is the difference between the measured lowest power and highest power for matings between multiple connectors and the same transceiver.

In order to reduce XPR between components, we must establish the sources of coupling variations. Once these are defined, it becomes an engineering problem to optimize component performance.

9.4.1. Optical Alignment

An optical parameter that plays an important role in coupling is the beam itself. Most beams can be characterized as approximately Gaussian: for example, a SM fiber's fundamental mode, an LD's far-field (although elliptical), the distribution of modes in a MM fiber, and the beam from an LED. Values for each of these parameters vary widely and often must be determined through direct measurement [15]. Rather than use separate terms for each (e.g., mode field diameter, spot size, numerical aperture, etc.), we will simply use source beam spot (SBS) and fiber beam spot (FBS), although the reader should be aware that different definitions of "width" are used: 1/e, $1/e^2$, full width at half maximum, and so on. When you are given a value for a beam spot, make sure you know which definition has been used. Even when all mechanical alignments are perfect, a difference in beam spots between components will reduce the coupling efficiency. Although the FBS is well measured and controlled, the SBS is usually poorly known or controlled.

9.4.2. Planar Alignment

Physical misalignments are the most obvious source of coupling variation when mating two optical components such as a ferrule and an OSA bore. We can define the interface plane as the plane normal to the optical axis. Longitudinal displacements occur along the (optical) z-axis and are not significant [11, 16]. We do consider some lateral misalignment in the *x* and *y* directions, or *r* (where $r^2 = x^2 + y^2$).

Ferruled fibers are inserted into bores in the OSAs that house the transmitter or receiver. Two style of bores are available: solid and split sleeve. A solid bore has complete rotational symmetry in the sense that it is a "perfect" cylinder as shown in Fig. 9.la. It is typically constructed of a rigid, ceramic material. The split-sleeve bore, in contrast, is separated lengthwise along the cylinder wall, as is shown in Fig. 9.lb. Its composition is of a more flexible material. The split-sleeve bore is designed so that its effective diameter will enlarge slightly to accommodate a larger ferrule.

The solid bore inner diameter (BD) and ferrule diameter (FD) also determine coupling performance. The difference in these two diameters creates an offset in solid bore TOSAs that can vary from insertion to insertion in the same module/connector pair. Hence, these two elements are critical in plug repeatability for solid bore TOSAs. Typical values for the BD are 2.5010–2.5025 mm. For the ferrule dimensions, some specifications require a diameter of 2.4992–2.5000 mm, within a 3-standard deviation criterion. See Fig. 9.2 as a reference. These parameters give the largest ferrule-to-bore difference possible of 3.3 μm. The minimum difference is 1 μm, and assuming normal distributions the average would be 2.15 μm. Beam centrality (BC) is the position of the laser beam axis with respect to the mechanical axis of the bore. It is desirable that the two be coaxial (making

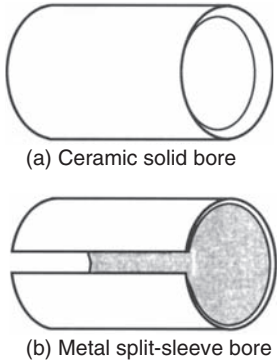


Figure 9.1 Solid and split-sleeve bores [17].

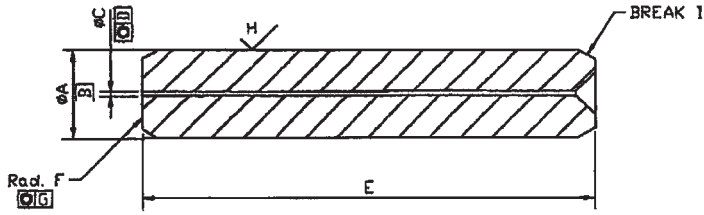


Figure 9.2 Example schematic of a ferrule.

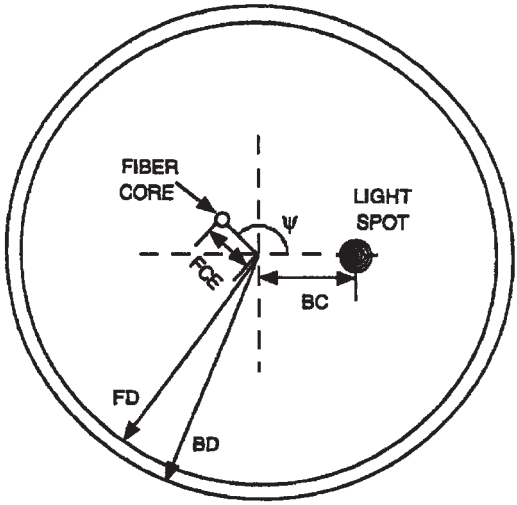


Figure 9.3 BC and FCE (exaggerated scale) [17]. *Note that the ferrule and bore centers coincide.

BC = 0) so that the laser beam center is exactly in the center of the bore. The ferrule/core eccentricity (FCE) is the corresponding centrality misalignment element in the connector. The FCE is typically controlled to <1.6 μm. It is nonzero when an imperfect ferruling process results in a fiber core that is nonconcentric with the ferrule.

These parameters are shown in Fig. 9.3. Note that to ease interpretation, it is assumed for now that the ferrule fits "perfectly" within the bore so that the ferrule and bore centers coincide. (The degree of "coincidence" is explained below.) Therefore, we can define an angle, Ψ, between the directions of BC and FCE with respect to some fixed coordinate system.(Ψ itself does not constitute a new "element.") For the BC and FCE given, the Ψ value will be randomly distributed throughout 360°C. For Ψ = 0″, the misalignments will have the least effect,

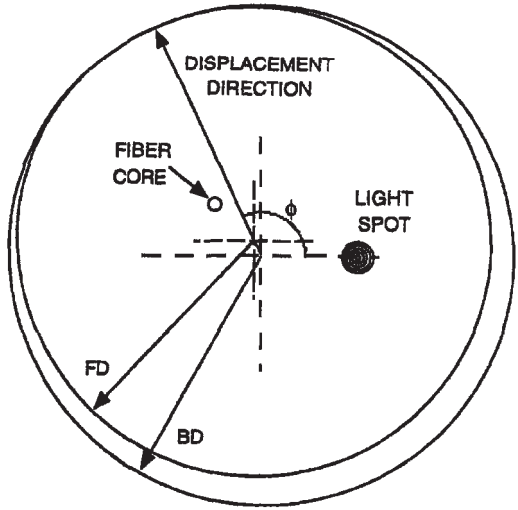


Figure 9.4 BD and FD (exaggerated scale) [17].

whereas $\Psi = 180''$ will place the fiber core the farthest from the light spot. The scales of the BC and FCE have been grossly enlarged in Fig. 9.3 for clarity.

These same elements are shown in Fig. 9.4 for an arbitrary direction, Φ, of displacement. If FD − BD = 0, then Φ is meaningless and we revert back to Fig. 9.3. Depending on the Ψ orientation between BC and FCE, Φ can either improve coupling or degrade it. Because Ψ and Φ are random variables, the best we could hope for is to establish a range of possible distances between the fiber core and the light spot.

The random fluctuation of all these dimensions can be reduced during manufacturing through a process known as tuning, whereby the eccentricity is intentionally aligned in a preferred direction [1].

It should be observed, however, that split-sleeve TOSAs will not be sensitive to these elements; by design, the split-sleeve walls flex to achieve a snug fit for a wider range of FD.

9.4.3. Angular Alignment

The ferrule/bore diameter difference can lead to a source of misalignment known as the angular offset, or tilt. However, due to the small lateral displacement of BD − FD = 3.3 μm, and the long length of the bore of approximately 4 mm, this tilt will be small. The largest tilt due to the ferrule/bore difference would be *0.05°*. This can be shown to contribute insignificantly to the CPR [16].

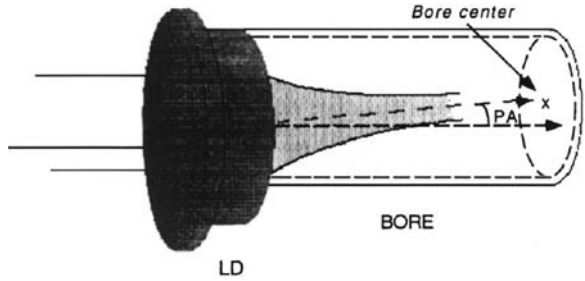


Figure 9.5 Laser PA [17].

There are, however, other possibly significant sources of tilt. One potential for tilted beams can be found in the TOSA. Even if the beam center spot can be successfully imaged to the bore center (i.e., BC = 0), the laser may possess some tilt as shown in Fig. 9.5. This tilt is known as the pointing angle (PA). As explained previously, before the laser and bore are welded, the laser is repositioned for maximum coupling. If there is a significant PA to the laser light, the weld will still be optimum for a small BC, as shown, and the engineer will be oblivious to the fact that any tilt is present!

The pointing angle can be as large as 3° and must be taken as an unknown. It is hoped that as datacom applications become more common, the need for well-controlled tolerances will drive manufacturers to produce modules with much smaller laser tilts. A necessary precursor to this optimization is the development of an accurate measurement technique.

The fiber end-face angle (EFA) is due to the polishing process of the ferruled fiber. It is common to intentionally polish fibers at angles ranging from 5 to 10″ in order to reduce back reflections from the fiber surface. The de facto industry standard is 8°.

However, even supposedly "flat" fibers may have a nonzero EFA. Fibers are typically arc-polished so that the hemispherical tip deforms slightly when two ferruled fibers are butt-coupled. This process, known as the physical contact polish, assumes that the polished apex coincides with the fiber core center. If the polishing tool is not precisely aligned, the end-face normal of the fiber will not be parallel to the fiber/ferrule axis. A nonzero fiber FCE further complicates this situation.

Refraction of the entering/emerging light will cause the effective light axis to be nonparallel, or tilted, with respect to the alignment axis. As seen in Fig. 9.6, a beam emerging from an obliquely polished fiber is refracted according to Snell's Law:

$$n_1 \sin(\text{EFA}) = n_0 \sin(\text{EFA} + \theta) \tag{9.1}$$

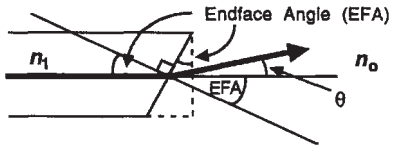


Figure 9.6 Fiber FEA.

It can easily be seen that when a fiber end face is polished at an angle EFA, the beam emerges at an angle θ with respect to the fiber axis (which is an angle EFA + θ with respect to the end-face normal). Using the small angle approximation, θ can be found:

$$n_1(\text{EFA}) \approx n_0(\text{EFA} + \theta)$$

or

$$\theta \approx \left(\frac{n_1}{n_0} - 1\right) \text{EFA} \tag{9.2}$$

Using 1.47 for the fiber's index of refraction, and a value of 1 for air, the relation between θ and EFA is θ ≈ (0.47) EFA. Note that for analytical purposes, θ can be seen as the fiber analog to the TOSA pointing angle.

Random sampling of nonrejected spherically polished fibers shows up to 0.3° of effective tilt, θ. The measurement of the end-face angle can typically be made using interferometric techniques so that samples with, for example, >l0.1° EFA can be rejected, but this will produce large numbers of unused cables, increasing manufacturing costs.

9.4.4. Module Alignment

The module itself can also contribute to coupling variations through the subassembly misalignment (SAM) and connector ferrule float (FF). SAM is the axial misalignment of the entire OSA/bore assembly from its desired true position within the transceiver module. Remember that we are discussing duplex interconnections so that both the transmitting and receiving bore/ferrule matings must be aligned. If one of the couples cannot be mated, then by design the other pair will be prevented from achieving successful plugging.

FF is the amount of lateral movement or float possessed by the connector ferrules. Too rigid a setting would impede ease of use and/or would make component damage more likely. Thus, spring forces built into the design allow a little "give." FF could compensate for a slight SAM. Of course, too large a FF could also prevent proper mating because of the "sloppiness" of the ferrule position. In split-sleeve OSAs, excessive SAM may cause a large off-axis force on the split sleeves that can cause plug repeatability problems or damage the split sleeve. It is impor-

tant to note, however, that these two elements do not strongly affect the plug repeatability or cross-plug range for solid bore systems because the bore stiffness overwhelms them. A detailed account of the forces involving FF and SAM is given in Ref. [2].

Other characteristics are the module shroud dimension (SD) and the connector body dimension (CD). The clearance between the two is the variable that affects plug repeatability and cross-plug range. Excessive clearance can create several mechanically stable plug modes that can affect plug repeatability in the same module/connector pair. Data indicate that this clearance has a much stronger effect on split-sleeve OSAs than on solid bores.

9.4.5. Putting It All Together

For reference, Table 9.1 provides typical values for the 12 elements. A typical CPR might be from −2.2 to −10 dBm-a 7.8-dBm range. All of the variables that affect the CPR are usually analyzed with a Monte Carlo technique to ensure that the CPR can be achieved. The CPR can be viewed as the sum of variations due to aging effects on the electrical components and to changes in coupling from mechanical/optical variations in the connectors and TOSAs.

Table 9.1

Typical Values of the 12 Elements.

		CPR Influence	
Element	Typical Value	Solid Bore	Split Sleeve
SBS	5–20 μm[a]	√√	√√
FBS	9.5 ± 0.5 μm for SM[a]	√	√
BC	3.3 ± 2.0 μm	√√	√√
FCE	<1.6 μm (0.7 μm)	√	√
BD	2.5010–2.5025 mm	√	
FD	2.4992–2.5000 mm	√	
PA	<3°	√√	√√
EFA	0.1–0.2° (barring rejection)	√	√
SAM	Sped. ± 0.34 mm		√
FF	Spec. ± 0.34 mm		√
SD	Spec. ± 0.125 mm		√
CD	Spec. ± 0.125 mm		√

Note: From Ref. [17]. √ determines the influence of a parameter's variability. If a parameter either has little variation or has little effect on the coupling efficiency to begin with, it will be ranked with fewer √. Given these measurements, it is desirable to predict how much variation to expect from a given product that could include any combination of values.

[a]Using $1/e^2$ width definition.

In general, all 12 elements affect XPR. However, if we limit ourselves to solid bore TOSAs, only 8 variables are operable. Four variables are related to the connector and 4 to the module. These variables are:

Module	Connector
BD	**FD**
BC	**FCE**
PA	**EFA**
SBS	**EBS**

Each of these factors will produce variability in both PR and XPR. The actual distance between the fiber core and the laser spot is essentially determined by BC, FCE, BD, and FD as shown in Fig. 9.7 (compare with Figs. 9.3 and 9.4). Knowing these quantities along with the directions of displacement Ψ and Φ, we can develop a relation for the distance R:

$$R^2 = \left[BC - FCE \times \cos\Psi - \frac{1}{2}(BD - FD) \times \cos\Phi \right]^2 + \left[FCE \times \sin\Psi + \frac{1}{2}(BD - FD) \times \sin\Phi \right]^2 \quad [9.3]$$

It is clear that the maximum value for *R* occurs when *@* and *4* are 180″ and the minimum occurs for both directions equal to zero. Thus,

$$R_{max} = BC + FCE + \frac{1}{2}(BD - FD)$$

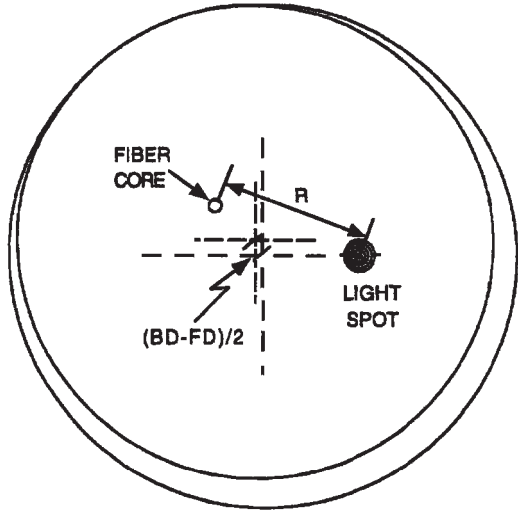


Figure 9.7 Fiber core and laser spot separation [17].

and

$$R_{min} = BC - FCE - \frac{1}{2}(BD - FD). \tag{9.4}$$

Clearly, BC, FCE, and DB – FD can give rise to an R_{min} of zero. Figure 9.8 shows the theoretical results for coupling from LDs with SBS = 9.5 and 14 μm to a standard fiber with FBS = 9.5 μm for purely radial offset *R*. Actual coupling losses can be up to 3 dB greater than those shown in Fig. 9.8, but the shapes of the experimental curves are surprisingly close to those in Fig. 9.8. It is believed that the difference is due to lens aberrations in the TOSA [7, 18].

Different modules will necessarily have different PAS, even if BC is well controlled. Figure 9.9 shows the effect on coupling loss that these various pointing

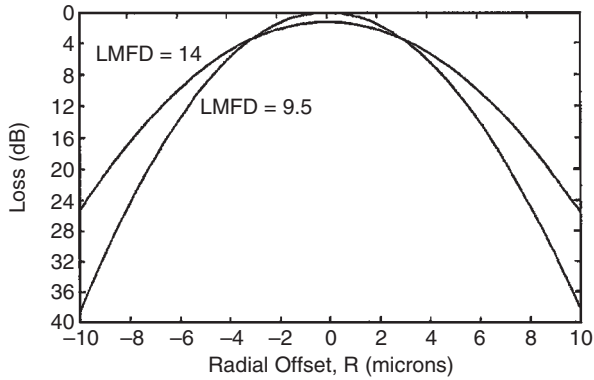


Figure 9.8 Coupling loss as a function of radial offset [17].

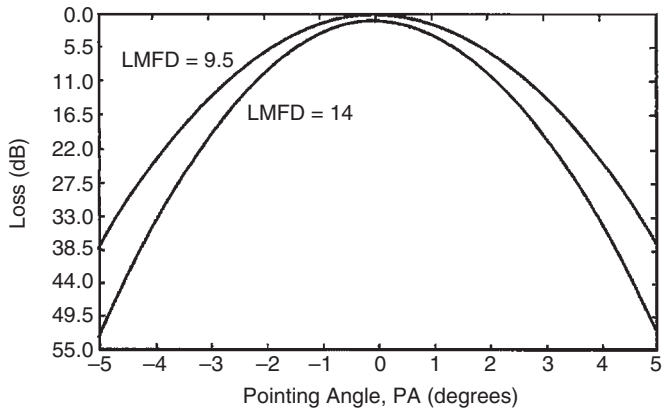


Figure 9.9 Coupling loss as a function of laser pointing angle [17].

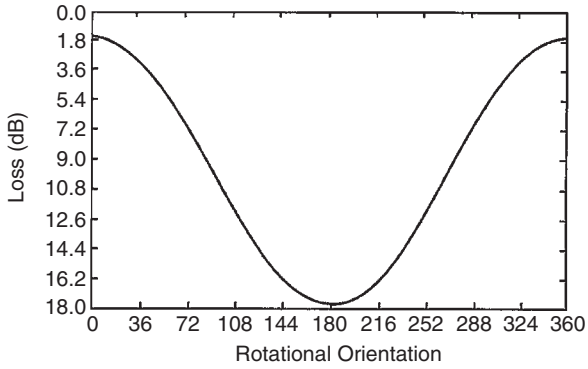


Figure 9.10 Change in coupling loss during fiber rotation for a large end-face angle [17].

angles will create. Generally speaking, this variation can be responsible for up to 10-dB differences in loss between matings of different modules with a given connector. However, plug repeatability will not likely be affected because the PA for a given module can be assumed constant.

The fiber EFA has a much less severe influence on coupling efficiency when it is assumed that other misalignments are small. In fact, because it is common that EFA < 0.2°, one can use Fig. 9.9 to estimate its effect on overall coupling loss, provided one sets PA = θ = 0.47 (EFA) from Eq. (9.2). The variability comes from different insertions of one connector/module pair. With 360° of possible rotation within the bore, a fiber with an EFA = 0.2° will produce minimal coupling variation. It should be noted, however, that many applications call for intentionally angled end faces to reduce backreflection into the source. The industry standard for the EFA in this case is 8°. Here, for the same parameter values as in the previous paragraph, different insertion orientations have dramatic effects, as shown in Fig. 9.10 [8].

In our discussions, we have primarily considered isolated misalignments: radial, axial, and angular. When all these are present, they can each become more significant to coupling. When the radial offset is significant, the PA has a more dramatic effect and vice versa. Nemoto and Makimoto [16] give analytic solutions when the three misalignment "categories" can be assumed to occur in the same plane. Out-of-plane misalignments are explored in Ref. [8].

9.5. ALIGNMENT TECHNIQUES

Alignment technology is critical to the production of low-cost optodatacom components. Some consider it to be the critical issue that may delay the proliferation of optical data communication. Currently, it can cost up to $100 to align and

secure a laser diode in a SM optical subassembly. Developing technologies to reduce this cost by one or two orders of magnitude is paramount to the proliferation of optodatacom into the consumer markets.

9.5.1. Alignment Requirements

Although SM alignment requirements are often stated as submicrometer, experience has shown that 1–3 μm control of the center of the laser spot to center of the SM fiber is usually adequate. MM technology is relaxed to approximately five times that of SM technology. Active (light source turned on), manual alignment technologies for manufacturing became available in the late 1990s to perform both SM and MM assembly. Unfortunately, these techniques are slow, serial, and hence expensive. Passive and active automated alignment techniques that may address this need are on the horizon.

9.5.2. Current Alignment and Securing Technology

As stated in previous chapters, a transceiver module generally consists of a ROSA and a TOSA assembled on a substrate with associated drive and mux/demux circuitry. The critical alignments that we are discussing are needed to assemble the TOSA. The alignment tolerances in assembling the ROSA are considerably more relaxed as discussed previously. To review the assembly process and structure of the TOSA, see Chapter *5*. In a typical TOSA, the optical axis of the laser (or other source) is first aligned to that of a lens. This alignment is currently performed with the light source active, using some type of robot. When the alignment is optimized, the robot measures a maximum in light output. At this point securing of the lens and laser positions is performed, usually using some type of welding or occasionally an adhesive. The welding process, though fast and strong, can also result in a shifting as the weld cools and shrinks. The assembled laser and lens are then aligned to the center of the bore of the TOSA. This process is also performed actively, aligning the light from the laser/lens combination with the bore center by measuring light coupled into a ferruled fiber in the bore. Using a ferruled fiber in the bore contributes a 1- or 2-μm offset immediately because there is clearance between the ferrule and bore and the fiber is never perfectly centered in the ferrule. These errors were discussed previously in this chapter: FCE and BD – FD. After the laser/lens combination is aligned by the robot to the bore/ferrule, another weld is formed to secure these two components. Unfortunately, another “weld shift” is usually experienced. Hence, with the weld shift and bore to ferrule error, it is typically difficult to obtain better than 3-μm alignment in a welded TOSA. These steps are time consuming and hence expensive, especially in SM technology.

9.5.3. Active Automated Alignment

Active automated alignment involves aligning a source to an optical detector (i.e., axis) while both source and detector are active. A primary advantage of active alignment techniques is that they produce known good assemblies with optimum or near-optimum alignment. This is in contrast to passive alignment techniques, which must be tested postmanufacture. Cost is believed to be the primary disadvantage with this alignment technique. The prevailing activity in optoelectronics packaging is to develop passive alignment technology for low-cost assembly. The most common method for active automated alignment involves mechanical alignment techniques. Recent developments in selective silicon etching and silicon micromachining have enabled development of combined active and passive alignment techniques, simplifying the assembly process.

Numerous automated alignment techniques have been reported in the literature, with each having unique advantage. A representative set can be found in Refs. [19–23]. A typical assembly process involves first the temporary insertion of an optical fiber into a large area photodetector to determine the maximum optical signal strength. This signal strength is used to determine the maximum relative signal strength during active alignment. Then the optoelectronic package is assembled, and the source and detector are activated. Next, after fiber is removed from the photodetector, it is inserted into the optoelectronic package and brought into the vicinity of the photodetector or source depending on the function of the package. Micromanipulation is then used to align the fiber to the photodetector or source. The end of the fiber is manipulated in five degrees of freedom (i.e., δx, δy, δz, $\delta\theta_x$, and $\delta\theta_y$) while the signal strength on the detector is monitored. The photodetector signal strength is maximized to find the optimum alignment. At this point, the fiber is permanently affixed in the package to secure the aligned position. Typical fixturing processes include soldering, adhesive attach, and welding.

Active automated alignment in optoelectronic assemblies is typically accomplished using one of three techniques for maximizing the detector signal as the fiber is precision aligned relative to the detector. Near-optimum alignment is achieved using simple linear five-axis optimization in which the fiber is moved independently along each of five axes (i.e., x, y, z, θ_x, and θ_y), to identify a local maximum signal strength [22]. Each axis is traversed independently, whereas previously analyzed axes are fixed in their local maxima positions. The primary drawback with this technique is that it rarely produces optimum alignment and is prone to locating local maxima having signal strength considerably lower than the maximum. A second active alignment technique utilizes minimum path search techniques to determine the optimum alignment. These search techniques are derived from fundamental research performed in robotic motion and

automated assembly. Minimum path searches involve incrementally moving the fiber along each degree of freedom (i.e., δx, δy, δz, $\delta\theta_x$, and $\delta\theta_y$) and measuring the signal strength. The path motion space is then analyzed to identify the incremental motion yielding the maximum positive gradient of the signal strength. The path variables are then incremented based on the local optima. The fiber is again incrementally moved along each degree of freedom, and the path of maximum response is again identified. This process is repeated until all combinations of incremental moves generate negative signal strength gradients corresponding to the optimum alignment. Identification of localized maxima is a potential problem with this technique, although various techniques have been developed to minimize the problem. With this technique, the fiber is incrementally moved along a path of increasing signal strength until the maximum is reached.

The third technique for active automated alignment is direct feedback control of the micropositioning system during active alignment [24]. One technique used for feedback control utilizes a phase-sensitive detector in the optoelectronic assembly and drives the fiber alignment axis using a sinusoidal dithering motion superimposed on a monotonic translation. The output of the phase-sensitive detector is directly proportional to the spatial offset error (used for feedback control) of the optical fiber from the optimum alignment position. Moreover, the polarity of the phase-sensitive detector output corresponds to the positive and negative approach directions for the optimum alignment position. For zero-offset positioning of the fiber from optimum (within 0.1–l%), an integral feedback controller can be used. Additional options include proportional-integral and proportional-integral-derivative as feedback control systems. The primary advantage of this technique is its rapid response and settling time for locating an optimum alignment. For multiaxis alignment, this technique has a significant speed advantage in that it allows optimization of the various axes simultaneously. This is done by having the micropositioning system actuators for the different axes dithered at different frequencies. Different phase-sensitive detectors are then used to extract the spatial offset information pertaining only to the appropriate axis. The spatial offset errors from each phase-sensitive detector are then used as the feedback control error signal for actuating the independent axis transducers controlling the fiber position.

Assembly of optical fiber is typically achieved using a microactuator to mechanically align the fiber relative to a detector or source. The manual alignment technique is the most common method. Here the fiber is grasped with a microgripper attached to a precision multiaxis stage and translated/rotated using precision lead screw actuators. To minimize vibration and backlash, the stages typically ride on air bearings and are mounted to large granite slabs. Actuator drive systems incorporate submicrometer encoders and closed-loop position control to ensure precise alignment.

9.5.4. Hybrid Active/Passive Automated Alignment

Due to the excessive cost associated with five degrees of freedom, precision actuators, and the relatively low-volume utilization of the precision assembly systems, recent efforts have focused on combined active and passive alignment assembly. Numerous techniques have been reported using various degrees of passive alignment [25–31]. Tewksbury et al. [32] also present a comprehensive review of optical interconnect technology. The primary advantage of combined active and passive assembly technology is significantly reduced cost. This is because combined active and passive alignment requires only one or two precision actuators compared with five for a full degree of freedom assembly. Numerous combined active and passive alignment techniques have been developed. We will focus on a few representative techniques here.

Due to the relatively high cost of optoelectronic component assembly and optical fiber alignment, there has been a heightened interest in fiber array assembly technology. Jackson et al. [33] present several approaches for aligning linear arrays of optical fibers (similar to ribbon cables used in electronics) to multichannel GaAs lasers and detector arrays. The assembly process is simplified using anisotropically etched precision silicon V grooves etched in a rectangular silicon block. The fibers are secured in the V grooves using high-temperature adhesives or solders forming a passively aligned subassembly. In the case of solder attach, the fibers must be metallized with a solder-wettable metallization system to ensure reliable adhesion and robust performance. This is particularly important for subsequent processing steps that can subject the assembly to numerous high-temperature thermal excursions. The fiber array and V-groove assembly significantly reduces assembly complexity by transforming the multifiber array into a single-fiber subassembly that can be aligned and assembled as a single unit. This enables multiple interconnects to be formed during a single assembly process and significantly reduces the cost per interconnect. To further simplify assembly of the fiber array to the vertical GaAs laser array or the detector array, the fiber V-groove subassembly is beveled on one end (typically at a 35 or 45° angle relative to the horizontal). The beveled end acts as a mirror reflecting the vertically transmitted light from the lasers into the fibers or from the fibers into the detectors. In the case of the 45° bevel, the fiber end must be metallized to enhance reflectivity and reduce coupling losses; the 35° bevel totally reflects the transmitted light and does not require metallization. The alignment of the fiber array to the laser array and/or detector array is achieved using standard precision actuator systems. To simplify the alignment process, photolithographically defined ridges and contacts are produced on the laser array and/or detector. These act as alignment faducials for precision alignment of the fiber array. This assembly can have an angular registration better than ±1° and positioning alignment within +3 μm [33]. As one would expect, these tolerances are well within the requirements for multimode assembly.

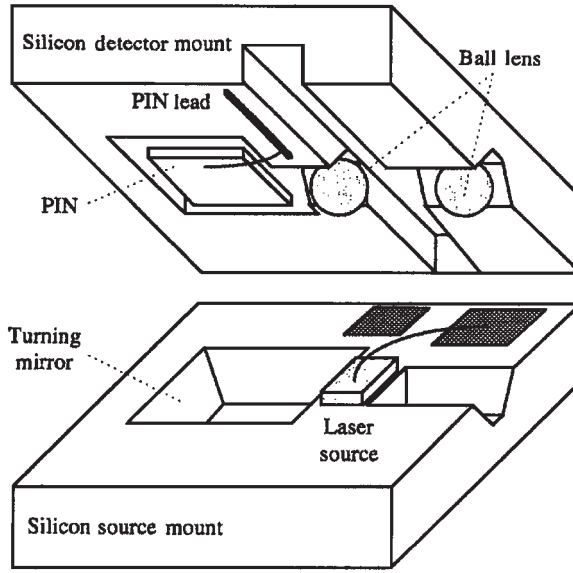


Figure 9.11 Schematic drawing of laser source and back facet monitor subassembly (adapted from Ref. [34]).

A second representative assembly processing using combined active and passive alignment is presented by MacDonald et al. [34]. The particular assembly integrates a laser, a back-facet monitor diode [positive-intrinsic-negative (PIN)], and coupling optics and is based on a silicon subassembly.

As shown in Fig. 9.11, the silicon source submount houses the laser. Registration of the active area of horizontal (side) emitting laser to the source submount is achieved with a solder column bond site as shown in Fig. 9.12. A large pad of solder is sputtered onto the source submount bond site, and textured compression bond sites are formed on stand-off columns of the source submount using selective silicon etching. The laser is then tacked, junction side down (i.e., the metallized side down), to the texture bond site. Mechanical alignment of the laser is achieved using a precision actuator system capable of compression bonding. Permanent bonding of the laser is achieved by reflowing the solder, which balls due to surface tension effects, forming a permanent electrical and mechanical bond. The process typically results in a ±2-μm registration of the laser active area to the source submount. The source and detector submounts also have selectively etched cavities and V grooves. One such element is the turning mirror in the source submount that reflects back transmitted light from the laser onto the surface-illuminated PIN on the detector submount. Key elements to the assembly are the front and back facet ball lenses, which are placed in selectivity etched lens holders (V grooves intersecting at 90°) that precisely establish the position

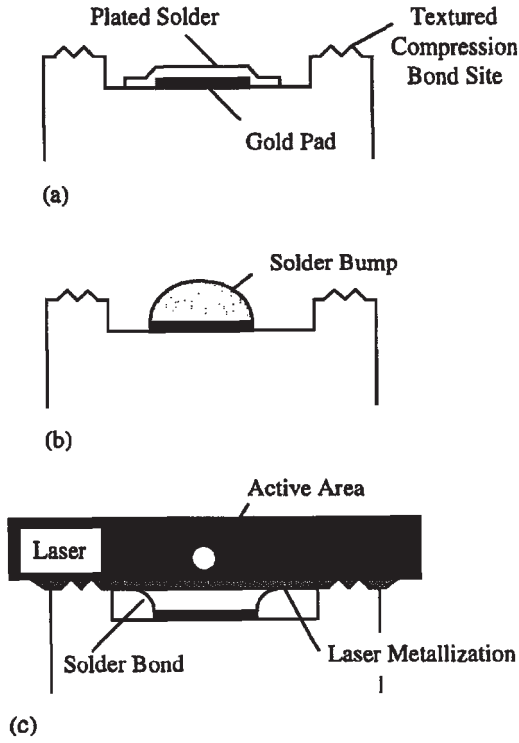


Figure 9.12 Solder column bond site featuring textured compression tacking and bump bonding (adapted from Ref. [34]). (a) As plated. (b) After solder reflow. (c) Laser tacked into textured surface and solder bump bonded.

of the spheres relative to the surface of the source and detector submounts. The position of the two lenses is controlled by photolithography and etching to within $\pm 1\,\mu m$. When the two submounts are assembled, the surfaces of the two submounts are brought into contact, aligning the laser and lenses in the direction normal to the submount surface (the z direction). The axial degree of freedom along the optical signal path is established by the internal references (V grooves and columns) defined by photolithography and selective etching. The lateral degree of freedom is established by sliding the two submounts relative to one another. Alignment is optimized along a single axis by actively monitoring the near-field intensity at the output of the front facet lens. Therefore, in this assembly technique, active alignment is used on only a single axis. The remaining axes are aligned passively.

A third combined active and passive alignment assembly process is presented by Solgaard et al. [35]. Their innovative optoelectronic packaging technology

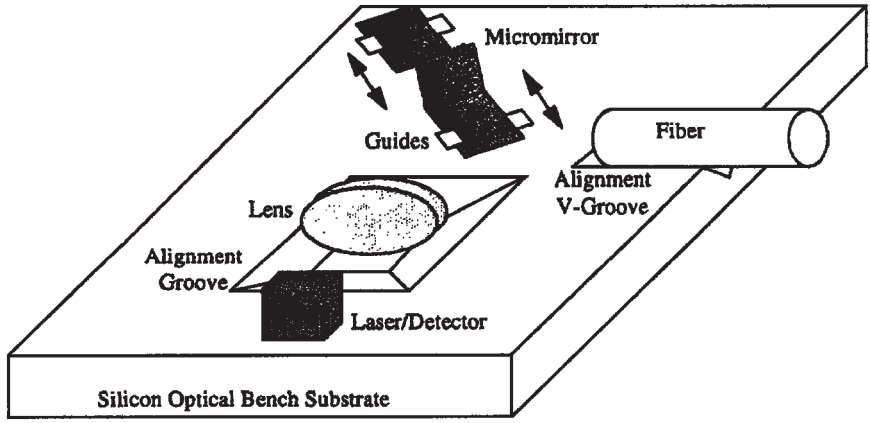


Figure 9.13 Fiber-coupled, semiconductor module with on-chip alignment mirror.

uses silicon surface micromachined alignment mirrors for active alignment of an optical fiber to a laser or detector. Their alignment technique is based on silicon optical bench technology in which active and passive optoelectronic components are assembled on a silicon chip. Typically, accurate passive positioning of the components is achieved by solder bumps and mechanical stops that are defined using conventional photolithography. Although this technology has found wide application, it does not provide adequate positioning accuracy for high-performance applications such as CATV laser modules. The technology Solgaard et al. have developed incorporate movable, micromechanical structures that are fabricated directly on the silicon chip by polysilicon surface micromachining [34]. The independent optoelectronic components are assembled to the package using standard optical-bench technology. The inherent inaccuracy of the passive alignment of the optoelectronic components is corrected by movable aligning microstructures as shown in Fig. 9.13.

Two schemes are used for alignment. The first uses an external servo feedback control system to automatically align the micromechanical structure. After alignment, the servo is disconnected and the micromechanical structure is left in a permanent position. The second technique is to integrate the servo feedback control, detector, and source in the package so that accurate alignment is guaranteed throughout the life of the package.

The alignment mirrors are micromachined polysilicon plated with gold to provide high reflectivity. The lower portion of the micromachined alignment mirror is connected by polysilicon hinges to another polysilicon plate that can translate linearly (i.e., a slider plate) along polysilicon guides etched in the silicon carrier. The top of the mirror is also hinged to a back support, which is hinged to a second slider. Moving the sliders simultaneously translates the mirror lin-

early. Translating one slider relative to the other tilts the mirror to a controlled angle. The two degrees of freedom (translation and tilt) provides the necessary beam alignment to compensate for the transversal offset of the lens and laser assembled using the conventional silicon optical-bench technology.

Potential concerns with this alignment technology include possible flexing of the structure during motion, dimensional tolerance loss due to loose hinges, and the possibility of slider motion during shock and vibration. However, due to the extremely light masses, inertial shock forces and vibration acceleration do not translate the micromirrors. This was demonstrated for shock tests up to 500 G. The micromachined structure is held in place by the friction forces in the slides and hinges.

9.5.5. Passive Automated Alignment

Passive automated alignment involves aligning a light source to an optic axis automatically without turning the light source on. This approach has a disadvantage in that it is not possible at the time of manufacture to determine if the alignment is successful because the light source is not turned on. However, with state-of-the-art manufacturing techniques any manufacturing loss due to this phenomenon can be minimized. One of the most promising technologies for passive automated alignment involves using the surface tension of solder. This technique is shown in Fig. 9.14. This self-alignment phenomenon was first observed at IBM [36]. Because their integrated circuit technology uses solder balls to interconnect, it was observed that if one did not precisely place the integrated circuit (IC) on the package, then the surface tension of the melted solder balls would ensure the centering of the IC. To use this technique to assemble the light source to a lens or fiber, the following process could be used. The light source is placed with an accuracy of approximately 25 μm by a standard IC die placement machine. Upon reflow of the solder, the surface tension moves the light source into alignment with the pads. Hence, the light source and fiber are aligned with approximately the same accuracy as the pads were to the optical axes. Lee et al. and Lin et al. [36–39] have shown that this technique is capable of accuracies down to micrometer levels.

With the advent of VCSELs, it is likely that a light source may be passively mounted directly to a fiber. A possible process to perform this assembly might be as follows. Photolithographic processes are used to form the pads (likely copper) on both the VCSEL and the fiber. Current art enables better than micrometer

Figure 9.14 Solder ball surface tension.

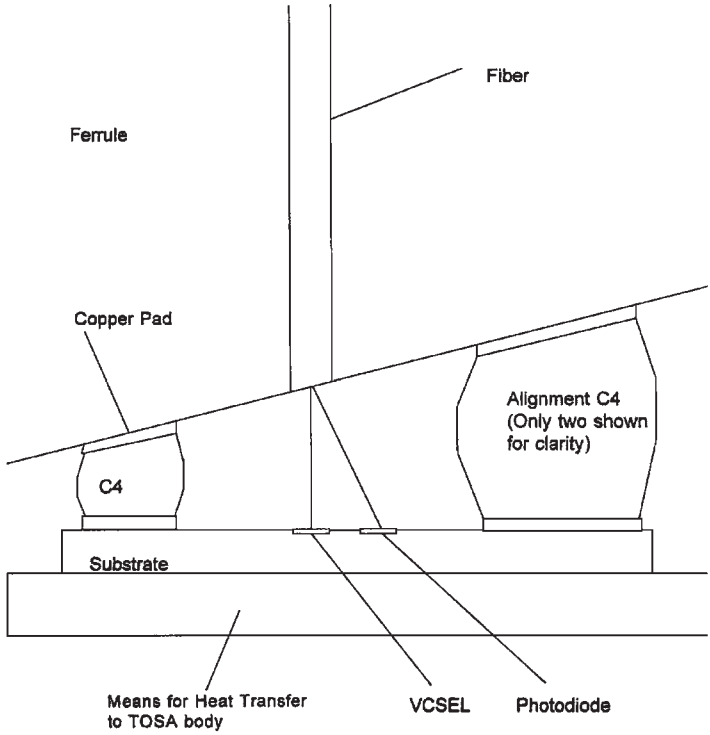


Figure 9.15 Passive alignment with solder balls.

location of the pads with respect to the VCSEL optic axis, in the VCSEL process. Unfortunately, a new tool would likely need to be invented to form the pads on the fiber to within micrometer accuracy with respect to the optic axis of the fiber. To form the pads accurately with respect to the optic axis of the fiber, the tool must focus on light emitted from the fiber and use this point as a reference to photolithographically image the pads in the correct spot. The lasers shown on this tool would image the pads in the photoresist on the fiber stub. To perform this operation, the lasers would have to be moved to the proper location. This movement would require state-of-the-art motion systems and good vibration damping to achieve the desired micrometer tolerance control. An additive or subtractive could be used to form the metal pads.

Because the SBS of the VCSEL can be controlled by design, it is possible to select it to match that of a SM fiber. In this case a lens is not needed to achieve reasonable coupling. Hence, a structure such as that in Fig. 9.15 would be possible. The fiber is tilted to minimize reflected light into the laser. It is evident that the solder balls require a gradation in size. This size gradation is possible with solder jetting technology [40].

Although somewhat more cumbersome, it is still possible to use this approach to align edge-emitting light sources to lenses or fibers, albeit with less coupling due to the typical mismatch between the SBS and the FBS.

To date, this process has still not been implemented into practice. Our belief is that the process could be implemented with initial costs of tens of dollars per SM alignment for process volumes of less than 100,000. As volumes approach a million, cost per alignment should approach several dollars or less. These kinds of price reductions will be needed to enable optoelectronic communication to become a consumer product.

9.6. CONCLUSION

Analyzing coupling efficiencies and the effects of alignment technology is an inherently complex process. This chapter has outlined the relevant parameters and the principles needed. However, for detailed calculations, many references have been presented that explore theoretical and experimental techniques. Currently, all models show a slight deviation in absolute coupling efficiencies, but the relative results are well established.

Various alignment methods have been presented, but it is understood that this area is still under development for optodatacom. Passive automated alignment of components is the ultimate goal of optodatacom manufacturing. Parallel components and VCSEL technology in particular present new opportunities to explore new manufacturing techniques and to reduce the cost of components.

REFERENCES

1. Vukovic, M. 1996, July. Analyze connector parameters to realize high performance. *Lightwave:* 50–56.
2. Paz, O., H. B. Schwartz, D. E. Smith, and E. B. Flint. 1992, December. Measurements and modeling of fiber optic connector pluggability. *IEEE Trans. Comp. Hybrids Manu*$ *Tech.* 15(6):983–991.
3. Aulet, N. R., D. W. Boerstler, G. DeMario, F. D. Ferraiolo, C. E. Hayward, C. D. Heath, A. L. Huffman, W. R. Kelly, G. W. Peterson, and D. J. Stigliani, Jr. 1992, July. IBM enterprise systems multimode fiber optic technology. *IBM J. Res.* Dev. 36(4):553–576.
4. Saleh, B. E. A., and M. C. Teich. 1991. *Fundamentals of photonics.* New York: John Wiley Sons.
5. International Electrotechnical Commission (IEC). International standard 825–1 and 825–2, safety of laser products. Central Bureau of the IEC, 3 rue de Varembe, Geneva, Switzerland.
6. American National Standards Institute. 1988. American National Standard for the safe use of optical fiber communication systems utilizing laser diode and LED sources. American National Standards Institute No. Z136.2.
7. Kawano, K., and 0. Mitomi. 1986, January. Coupling characteristics of laser diode to multimode fiber using separate lens methods. *Appl. Opt.* 15(1):136–141.
8. Clement, D. P., and U. Osterberg. 1995. Laser diode to single-mode coupling using an “out-of-plane misalignment” model. *Opt. Eng.* 34(1):63–74.

9. Hudson, M. C. 1974 (May). Calculation of the maximum optical coupling efficiency in multimode optical waveguides. *Appl. Opt.* 13(5): 1029–1033.
10. Kogelnik, H. 1964. Coupling and conversion coefficients for optical modes. Proceedings of the Symposium on Quasi-Optics, June 8–10. In *Microwave research institute symposia series,* Vol. XIV. Brooklyn, N.Y.: Polytechnic Press of the Polytechnic Institute of Brooklyn.
11. Marcuse, D. 1977, May/June. Loss analysis of single-mode fiber splices. *Bell Syst. Tech. J.* 56(5):703–718.
12. Miller, C. M. 1986. *Optical fiber splices and connectors.* New York: Dekker.
13. Neumann, E. G. 1988. *Single-mode cables.* Heidelberg: Springer-Verlag.
14. Snyder, A. W., and J. D. Love. 1983. *Optical waveguide theory,* Part 1. London: Chapman & Hall.
15. Anderson, W. T., and D. L. Philen. 1983, March. Spot size measurements for single-mode cables-A comparison of four techniques. *J. Lightwave Tech.* 1(1):20–26.
16. Nemoto, S., and T. Makimoto. 1974. Analysis of splice loss in single-mode fibres using a Gaussian field approximation. *Opt. Quantum Electron.* 11:447–457.
17. Lasky, R. C., U. Osterberg, and D. J. Stigliani, eds. 1995. *Optoelectronics for datacommunication.* New York: Academic Press.
18. Karstensen, H., and K. Drogemuller. 1996, May. Loss analysis of singlemode fiber couples with glass spheres or silicon plano-convex lenses. *J. Lightwave Tech.* 8(5):739–747.
19. Bargar, D. S. 1988. An automated fiber alignment, fixing, and hermetic sealing system. *SPIE* 994:11–17.
20. Bristow, J., Y. Liu, T. Marta, K. Johnson, B. Hanzal, A. Peczalski, S. Bounnak, Y. S. Liu, and H. Cole. 1996. MCM board level optical interconnects using passive polymer waveguides with hybrid optical and electrical multichip module packaging. *SPIE* 2691:18–24.
21. Gabler, C., L. Li, S. Hackwood, and G. Beni. 1986. An optical alignment robot system. *SPIE* 703:8–28.
22. Goldman, L. S. 1972, May 15–17. Proceedings of the 22nd Electronic Components Conference, 332–339, Washington, D.C.
23. Karioja, P., K. Tukkiniemi, V. Hikkinen, and I. Kaisto. 1993. Inexpensive packaging techniques of fiber pigtailed laser diodes. *SPIE* 1851:48–53.
24. Goodwin, J. 1986. Dynamic alignment of small optical components. *SPIE* 703:2–7.
25. Han, H., J. E. Schamm, J. Mathews, and R. A. Boudreau.*SPIE* 2691:118–123.
26. Kalman, R. E., E. R. Silva, and D. E Knapp. 1996. Single-mode array optoelectronic packaging based on actively aligned optical waveguides. *SPIE* 2691:124–129.
27. Kilcolyne, M. K., J. W. Scott, and J. Plombon. 1993. Packaging of optical interconnect arrays for optical signal processing and computing. *SPIE* 1851:80–88.
28. Schmid, P., and H. Melchior. 1984. Coplaner flip-chip mounting technique for picosecond devices. *Rev. Sci. Instrum.* 55:1854–1858.
29. Spitzer, M. B., D. P. Vu, and R. P. Gale. 1996. Applications of circuit transfer technology to displays and optoelectronic devices. *SPIE* 269133442.
30. Weber, R., E. Fidorra, M. Hamacher, H. Heidrich, and G. Jacumelt. 1993. Multi fiberkhip coupling and optoelectronic integrated circuit packaging based on flip chip techniques. *SPIE* 1894:149–152.
31. Yap, D., W. W. Ng, D. M. Bohmeyer, H. P. Hsu, H. W. Yen, M. J. Tabase, A. J. Negri, J. Mehr, C. A. Armiento, and P. 0. Haugsjaa. 1996. RF optoelectronic transmitter and receiver arrays on silicon waferboards. *SPIE* 2691:110–117.
32. Tewksbury, S. K., L. A. Hornak, H. E. Nariman, and S. M. Lansjoen. 1993. Opportunities and issues for optical interconnects in microelectronic systems. In *Optoelectronics: Technologies and applications,* eds. A. Selvarajan, K. Shenai, and V. K. Tripathi, Bellingham, Wash.: SPIE. 703 2–7.

33. Jackson, P., A. J. Moll, E. B. Flint, and M. F. Cina. 1988. Optical fiber coupling approaches for multi-channel laser and detector arrays. *SPIE* 994:40–47.
34. MacDonald, W. M., R. E. Fanucci, and G. E. Blonder. 1993. Si-based laser sub-assembly for telecommunications. *SPIE* 1851:42–47.
35. Solgaard, O., M. Daneman, N. C. Tien, A. Friedberger, R. S. Muller, and K. Y. Lau. 1995. Optoelectronic packaging using silicon surface-micromachined alignment mirrors. *IEEE Photon. Tech. Lett.* 7:41–43.
36. Lin, W., S. K. Patra, and Y. C. Lee. 1995. Design of solder joints for self-aligned optoelectronic assemblies. *IEEE Trans. Comp. Packaging Manuf. Tech. Part B,* 18:543–551.
37. McCroarty, J., B. Yost, P. Borgesen, and C. Y. Li. *Mater. Res. Soc. Symp. Proc.* 264:423–435.
38. Patra, S. K., J. Ma, V. Ozguz, and S. H. Lee. 1994, *SPIE* 2153:118–131.
39. Patra, S. K., and Y. C. Lee. 1991. *Electron. Packaging* 113:337–342.
40. Snyder, M. D., and R. Lasky. 1995. Proceedings of the MRS, Spring 1995 meeting. San Francisco.

10

Packaging Assembly Techniques

Glenn Raskin
Motorola, Incorporated, Chandler, Arizona

10.1. PACKAGING ASSEMBLY—OVERVIEW

Packaging can be thought of on many levels and serves many different functions. Its basic function is to transfer energy from one level to another. Whether that energy is electrical, thermal, or optical defines the level or type of packaging function. Traditionally, one considers the first level of packaging as that which deals with the interconnection of semiconductor or optoelectronic devices to some kind of leadframe. The second level of packaging concerns itself with interconnection of the aforementioned leadframe to a circuit board, on which other packaged devices are also mounted.

10.1.1. Requirements for Optoelectronic Packaging

The requirements for packaging drive the technologies that are used and the manufacturing techniques that are incorporated. In the area of optoelectronics, one is primarily concerned with the transfer of optical and electrical energy from board level to the device. For semiconductor devices the focus is typically limited to the electrical domain. Other areas of concern and interest in packaging focus on thermal transfer, mechanical strength, and electromagnetic and other considerations. The discussions in Sections 10.2, 10.3, and 10.4 apply to optoelectronic as well as microelectronic assemblies.

Before focusing on the specific details, it is useful to look at the overall picture. Some of the key considerations in packaging assembly include electrical, optical, and thermal performance of the package; reliability, cost, manufacturability, testability, protection, mechanical integrity, size, and weight. In any given application, one must consider these factors and design a package that meets all of the requirements. It may be necessary to make compromises and tradeoffs in each of

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

the areas when making packaging decisions. For optoelectronics the requirements that drive device packaging revolve around some of the key characteristics of the system itself. If one is trying to construct a single-mode package, the alignment tolerances to a single-mode fiber are much greater than those of a multimode system. These tighter alignments may require certain packaging techniques that limit the user's choice of technologies. Another key input required for optoelectronic packaging is the amount of acceptable attenuation, or coupling loss, that is allowed. If the power budget allocated for the packaging is not great, efficient coupling becomes extremely critical.

Another key consideration in optoelectronic packaging is the thermal and reliability requirements. A laser's performance and expected lifetime are much more sensitive to its operating temperature compared to classical integrated circuits. This strong dependence of lifetime and reliability on temperature has limited the applications of laser devices and driven the development of more exotic thermal packages for optoelectronics.

Finally, one of the most important considerations, and the one typically given the least ink in text, is that of cost and manufacturability. In the end, whatever technology and performance is used, the product must be cost effective to the end user. If the required technology to meet the performance needs of the product is costly, often another avenue must be taken. Because cost is a prime driver in the industry, we must frequently compromise in other areas of performance to meet the cost goals of the product itself.

10.2. FIRST-LEVEL INTERCONNECTS

First-level interconnects are those that connect the device to the leadframe. A variety of interconnect technologies are used in manufacturing; this text will review the three most common types: wire bonding, tape automated bonding (TAB), and flip chip.

10.2.1. Wire Bond

The most common of all first-level interconnects in the microelectronic as well as the optoelectronic industry is known as wire bonding. Wire bonding is done using a number of techniques and in a variety of packages.

The three most common wire bonding processes are thermocompression, thermosonic, and ultrasonic.

10.2.1.1. Thermocompression and Thermosonic Wire Bond

Thermocompression bonding was the earliest type of wire bonding, demonstrated by AT&T in 1957 [l]. Thermocompression bonding is typically done using ultrapure (>99%) gold wire. A wire bonder is composed of a stage that holds the

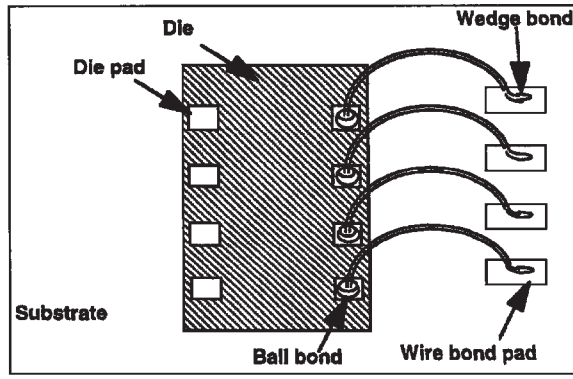


Figure 10.1 Schematic showing ball and wedge wire bonds.

package and the bonding head, which feeds and places the wires. The bonding head forces the gold wire into contact with the die or substrate metallurgy that forms a bond through heating. Typical bond head temperatures between 300°C and 400°C are required to form the bond.

Thermosonic bonding is very similar to thermocompression bonding except that the wedge bond is performed using ultrasonic excitation to assist in bond formation. This permits the use of lower temperatures than in thermocompression bonding.

Thermocompression and thermosonic bonding typically form a ball bond when contacting the device and a wedge bond onto the substrate. A schematic of ball and wedge thermosonic bonds is shown in Fig. 10.1. The process flow for both thermocompression and thermosonic wire bonding is the same. In either case we begin with a capillary holding a gold wire, typically from 0.001 to 0.05 in. in diameter, near its end. The wire that extends beyond the capillary is heated using a hydrogen flame or through capacitive discharge. This forms a ball on the end of the wire two or three times the diameter of the wire itself. This ball is then compressed, and the wire is heated to form the bond on the die (the substrate is also typically heated to a background temperature). The wire is then drawn through and looped over the die edge to the substrate. A wedge bond is then formed on the substrate, the wire is severed, and the wire is fed through the capillary in preparation of the next bond [2].

10.2.1.2. Ultrasonic Wire Bond

Ultrasonic bonding differs from thermocompression bonding in that it forms only wedge bonds. Industry-standard ultrasonic bonding utilizes aluminum wire and can be accomplished on very narrow pad pitches. With ultrasonic bonding the wire is typically fed through a Sono-trode or stylus. The stylus forces the wire

against the bond pad and applies a burst of ultrasonic energy to the interface. This process forms a cold wedge between the wire and metal on the bond surface. The stylus then loops up above the die surface and extends to the next bond point for the next wire bond.

10.2.1.3. Die Attach for Wire Bond

In order to wire bond to a device, the bottom of the device must be attached to a substrate of some kind, as shown in Fig. 10.1. The process of die attach can be accomplished utilizing a number of different materials from solders to organic adhesives to filled glass compounds [3]. Metallurgical attachment of the die to the substrate can be accomplished by metallizing the backside of the wafers. The die is then attached to the substrate using solder alloys. Solders offer low thermal and electrical resistance and are typically alloys of gold or lead.

Organic adhesives, most often epoxies, are the most common form of die attach. The adhesives are filled with precious metals to aid in thermal and electrical transfer. Epoxies provide reasonable performance and have been developed for highly manufacturable and relatively low-cost implementations. Glass adhesives, filled with silver, are a recent development that have not achieved wide acceptance in manufacturing areas.

10.2.1.4. Wire Bond Reliability

Wire bonding is a common process that has achieved very high yields and reliability levels for most applications. Many of the manufacturing defects have been eliminated as wire bond technology has reached its current maturity level. The most common reliability issues encountered in wire bonding today are due to gold aluminum mixing, die attach voids, and stress-induced creep. The mixing of gold and aluminum can lead to Kirkendahl voiding in extreme cases. This can be avoided by limiting the amount of time a unit is exposed to extreme (>400°C) temperature, by limiting the amount of impurities in the gold, and by forming the correct alloys in the bond. Stress-induced creep can occur in metallic joints due to low-melting eutectics (e.g., thallium) that weaken the grain boundaries. Once again, purity in the gold wires is key to avoiding reliability problems [4].

Reliability issues are also a concern in the die attach process. If excessive voids or bubbles are present in die attach material, failure can occur with repeated thermal cycling. Failures in this case can be loss of bond line, leading to poor thermal conduction from the die to the substrate. Stresses can be so severe in die attach defects that cracking of the die itself occurs.

10.2.2. TAB

Tape automated bonding was developed in the 1960s by General Electric and is in large use today for devices that require low profiles (e.g., display line drivers) or very high lead counts. TAB differs from wire bonding in that the leadframe

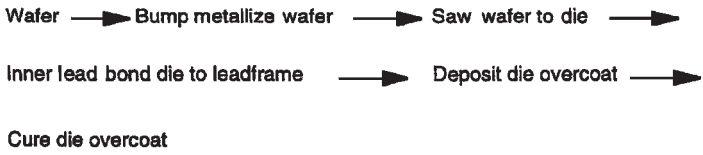


Figure 10.2 TAB flowchart.

in this embodiment is directly attached to the device. The flowchart in Fig. 10.2 describes a typical process flow for assembling TAB devices. TAB interconnection requires additional wafer processing to add a bump structure. Bump processing as well as the technology of interconnection will be discussed in this chapter.

10.2.2.1. TAB Bump Structures

Tab bump structures typically contain an adhesion layer, a barrier metal, and a plated bump. An adhesion layer helps promote adhesion between the underlying wafer metallization (typically an aluminum alloy) and the bump metal (typically gold). The composition of the layer can be copper, titanium, or chromium among others. The barrier metal serves the purpose of separating the bump metallization from the underlying metal. As previously discussed, one must be concerned with gold aluminum alloys. The barrier metal prevents intermixing of the aluminum and gold metallurgy.

Barrier metals are typically sputtered across the whole wafer and are composed of any number of alloys, including copper, palladium, platinum, or nickel [5]. After sputtering, the wafers use a mask-defined photoresist to determine where the gold bumps will be located. The bumps are plated in an electrolytic bath. Following the plating process, the photoresist and unwanted adhesion/barrier metallization are striped off the wafer.

10.2.2.2. TAB Tape Leadframes

The TAB leadframe is typically made of copper, is surface plated with either gold or tin, and comes in numerous formats. The leads are typically supported on a polyimide film that can be chemically etched, mechanically punched, or laser etched to provide the necessary windows where the leads need to be exposed.

High-performance tapes have also recently been implemented that use two metal layers. The additional metal layer can be constructed to help in power/ground distribution, thereby improving the electrical and thermal performance of the leadframe.

10.2.2.3. TAB Bonding

Attachment of the TAB leadframe to the die is commonly referred to as the inner lead bond. The leadframe is attached to the substrate in the area known as the outer lead bond. The bond is usually a single metallurgy gold bond or eutectic gold tin bond. The ILB process is performed using either thermocompression or thermosonic single-point bonding. The most common process used today is gang thermocompression bonding. The advantage of a gang process, as opposed to wire bond or a single-point TAB attach, is in throughput. All the interconnections in a gang process are made simultaneously, in a single operation, allowing for excellent throughput. The disadvantages of gang bonding include the requirement of high levels of planarity, alignment, and uniformity across the interface. Single-point ultrasonic TAB bonding is similar to wire bonding.

Gang thermocompression bonding requires a heated tool, better known as a thermode, to join the die to the leadframe. The thermode, whether a solid block or multiple pieces, is heated to a temperature slightly above that required to make the metallurgical bond (approximately 300–400°C for tin-gold and 400–500°C for gold-gold bonding). The compression force works to guarantee good thermal contact between the thermode, leads, and the bumps. The force applied typically ranges from 10 to 50 g per lead. The dynamics of the process itself are very different when comparing a eutectic bond to a gold-gold thermocompression bond. Tin-gold eutectic bonding forms a low-melting-point liquidous alloy at the lead bump interface. This alloy allows for a lower stress, lower temperature bond compared to the single-metallurgy bond.

Akin to wire bonding, thermosonic bonding techniques can also allow for lower bonding temperatures by using ultrasonic energy to aid in the attachment of the leadframe to the die. Thermosonic bonding is done in a single-point configuration and therefore does not have the throughput advantages of thermocompression gang bonding.

10.2.2.4. TAB Tape Reliability

TAB interconnections have shown themselves to be extremely reliable in a number of applications. As with wire bonding, care must be taken in the manufacturing process to guarantee good reliability. Key areas in ensuring reliability are in the integrity of the barrier/adhesion metals. The actual bond reliability relies on a high level of purity in the metallurgical components and especially in surface quality. As with most assembly processes, the time the device is exposed to elevated temperatures must be limited. This is especially true for many optoelectronic devices in which low-temperature films are often used to limit the stress and interdiffusion of the laser die.

10.2.3. Flip Chip

Flip-chip interconnection relies on direct attachment of the device to the substrate. Wire bond and TAB interconnects rely on peripheral attachment. In other words, the interconnection is made on the outside edges of the die. Flip chip, on the other hand, can be both a peripheral and an area array attach. By using the area array approach, a much higher density of interconnections can be achieved. This is shown in Fig. 10.3, in which the possible number of pads for typical wire bond pitches of 5.6 and 4.8 mils is compared to that of a typical flip-chip array pitch of 10 mils over a range of die sizes. For designs in which the number of interconnections are large, this can offer significant electrical and cost advantages. Another advantage of flip-chip attachment is that it can provide perhaps the lowest inductance interconnect between the die and the substrate. Low inductance is achieved by minimizing the length of the interconnect itself. Typical bump heights are on the order of 0.004 in. compared to 0.100 in. for wire bonds and 0.050–0.100 in. for TAB leads.

There are three major methods for flip-chip attach; C4 (controlled collapse chip connection), plated solder, and z-axis films. The most common embodiment is C4, whose history is closely tied with developments at IBM in the early 1960s [4]. In this embodiment an evaporated tin lead solder bump is used as the interconnect media between the die and the substrate. The plated solder solution is similar to the C4 process but relies on a plated bump rather than on an evaporative deposition. A typical flowchart that details the C4/plated bump processes is shown in Fig. 10.4.

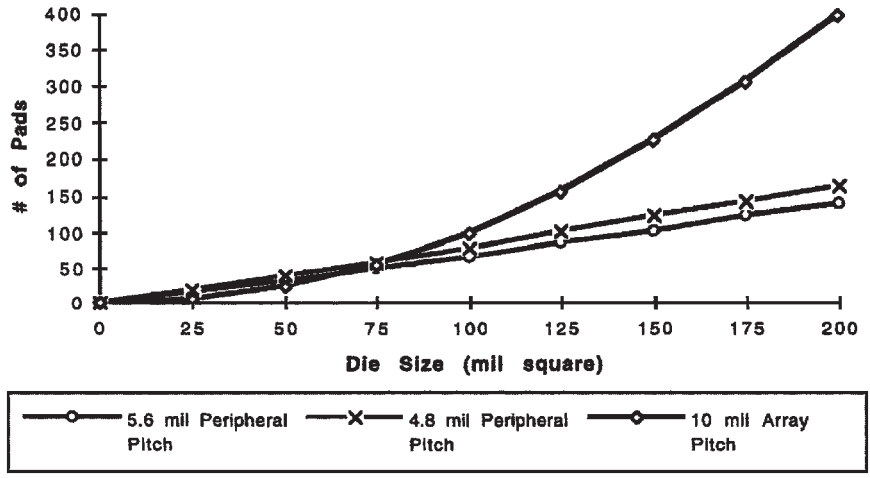


Figure 10.3 Comparison of the number of interconnections with peripheral and array technologies.

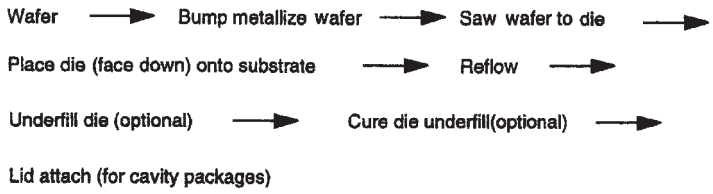


Figure 10.4 Flow diagram for C4/plated bump processes.

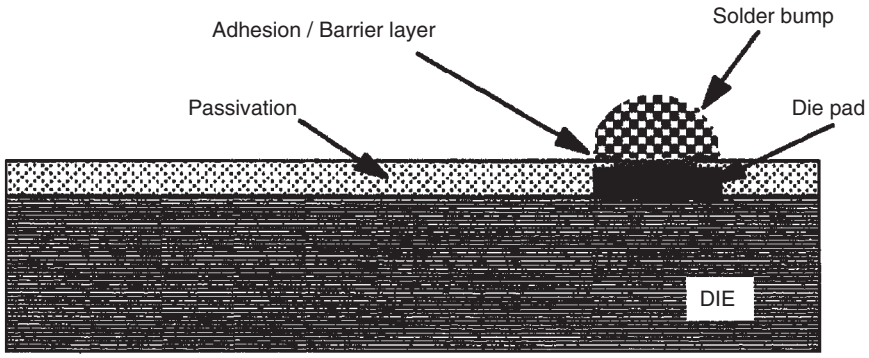


Figure 10.5 C4 bump cross section.

The z-axis interconnects are a relatively new development that rely on adhesive films filled with conductive particles. In compression the conductive particles form a low-resistance path-making contact to the substrate. Though promising, the use of conductive adhesives is currently very limited in the industry.

10.2.3.1. Flip-Chip Bump Processes

The bump process metallurgy and structure can be readily compared to the bump process used in the TAB system. Whereas the TAB system is usually reliant on a gold bump system, flip chip relies on tin lead solder metallurgy.

C4 bump processing requires numerous steps to achieve the structure shown in Fig. 10.5. Wafers with vias in the top passivation over the pads are clamped to metal masks. After clamping the mask directly onto the wafer the ball-limiting metallurgy (BLM) is evaporated. Typically, this is a film of chromium, copper, and gold deposited in a single evaporator. The sandwich of metals that make up the BLM serves the purpose of adhesion and barrier layers as well as oxidation protection prior to solder deposition. The solder is then deposited typically in a separate evaporator. The metal mask is removed, and the wafer is then reflowed.

Reflow of the bumps is accomplished in an inert environment (usually H2 at 350°C) that transforms the bumps into the spherical shape as shown in Fig. 10.5. The spherical shape is formed during reflow due to the surface tension of the solder bump. This surface tension is an important aspect of solder-attach processes because it provides for what is commonly known as self-alignment. Upon reflow, the liquidous metal will move to the lowest energy state it can achieve. This low-energy state centers the solder ball on the BLM [7].

Solder-plated bumps are manufactured in a very similar fashion to TAB-plated bumps. The only difference in the process is that solder is plated onto the wafers rather than onto gold. Numerous solder alloys have been considered. The most common for a solder bump is 95% lead/5% tin. This provides a high melting point (approximately 315°C) that can withstand the temperature that the package is exposed to during its attachment to the lower temperature solders (such as eutectic 63% tin/37% lead with a melting point of 183°C).

10.2.3.2. Flip-Chip Substrates

Solder flip-chip attach processes are very much dependent on the substrates to which the die is being coupled. Ceramic substrates are usually metallized with copper or nickel films. Organic substrates, such as printed circuit boards, use lower melting point solders that allow for joining at lower temperatures. The footprint of the substrate is a mirror image of the die. This is shown schematically in Fig. 10.6 [8,9].

10.2.3.3. Flip-Chip Attach

Attachment of the solder-bumped die to the substrate is a fairly straightforward process. Flux is applied to the substrates to enhance attachment, and the die is

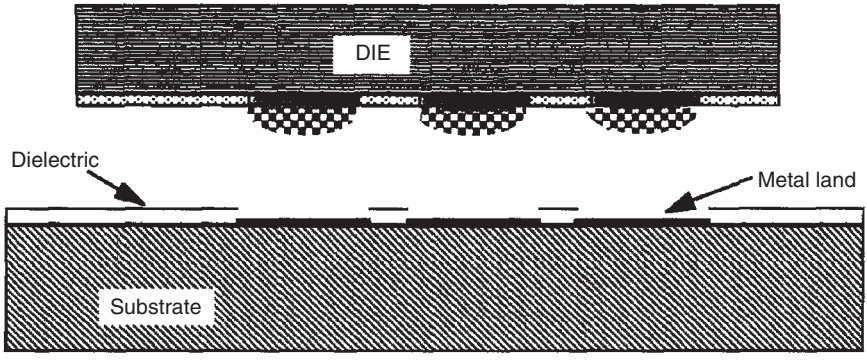


Figure 10.6 Flip-chip die attached to a substrate.

placed fairly accurately onto the substrate, as shown in Fig. 10.6. The assembly is then reflowed, and the connection of the die to the substrate is made in one process. In this way, hundreds of interconnections can be made quite rapidly. The placement of the die onto the substrate must be good enough so that some of the bumps are on the pad, but it does not need to be precise. Upon reflow, the self-alignment of the solder helps to correct for small misalignments. After attachment, the residual flux is cleaned using any of a number of solvents or water. At this point the package may be capped and assembly completed. In some applications (high bump counts, large die, and large thermal expansion mismatches between the substrate and die), an underfill is applied between the die face and the substrate. The reasons underfill is applied will be covered in more detail in the following section. The underfill itself is an epoxy that is dispensed around the die edge. Surface tension pulls the underfill into the area between the die and substrate. Following this process an underfill cure is required.

10.2.3.4. Flip-Chip Reliability

Because the die are being attached directly to the substrates, the amount of thermal expansion mismatch between the two is critical. The stress caused by the mismatch is mostly absorbed in the bumps themselves. Thermal expansion mismatch can lead to excessive stress on the bumps and reliability failures if not properly managed. For alumina ceramic substrates, the coefficient of thermal expansion (CTE) is approximately 6×10^{-6}/°C, matching silicon and gallium arsenide fairly well. For organic substrates, the CTE is much higher (approximately 12 to 16×10^{-6}/°C), and the mismatch stresses to the die are higher [10]. These mismatches cause a number of considerations in the application of flip-chip attachments. The first consideration is to use only materials that are well matched (such as alumina and silicon). The second consideration is that one may limit the size of the field of bumps that is allowed for flip-chip die. By limiting the bump field's size, the amount of stress on the outer bumps, where the extreme of the expansion mismatch is observed, is lessened. The third way to reduce the stress of thermal mismatch is to use an underfill epoxy between the die and the substrate. By acting as a compliant interstitial member, the epoxy reduces the effective stress that is seen by the bumps themselves.

Long-term effects of solder creep and fatigue are well known and can be avoided through careful design and the matching of materials as discussed previously. If one uses these guidelines, the reliability of solder-bumped flip-chip assemblies is excellent.

10.3. PACKAGE TYPES

Numerous kinds of packages and formats are available in the industry. The form factor of a package and its material makeup are driven by the application in which it will be used. Packages provide mechanical and environmental protec-

tion to the device while providing the electrical interconnections between the device and the circuit board. Packages in the optoelectronic industry can be divided into three categories; metal, plastic, and ceramic.

10.3.1. Metal Packages

Metal packages were the first packages used for the earliest transistors. As the microelectronics industry moved toward higher levels of integration, other packaging technologies were developed. The ceramic and plastic technologies were driven by the higher integration levels and the need for greater interconnection density between the device and the package. Metal packages in the form of transistor outline (better known as TO cans) are used widely in the optoelectronic industry today as the basis for numerous devices. These packages provide a cost-efficient method to package devices within an optoelectronic package. Schematic drawings of a TO can are shown in Fig. 10.7 [11].

TO cans utilize wire bonds (as discussed in Section 10.2.1) to interconnect the device to the package. As can be seen from Fig. 10.7b, the interconnection is made to the base of the package and to a post connected to the leads. TO cans are often integrated into optoelectronic modules in which other components (integrated circuits, discretes, etc.) are present. This provides the end user with a compact subassembly that serves a number of uses in optoelectronic products.

10.3.2. Plastic Packages

In a plastic package, the die is typically attached to a leadframe with any of the interconnection technologies discussed in Section 2. The leadframe and device

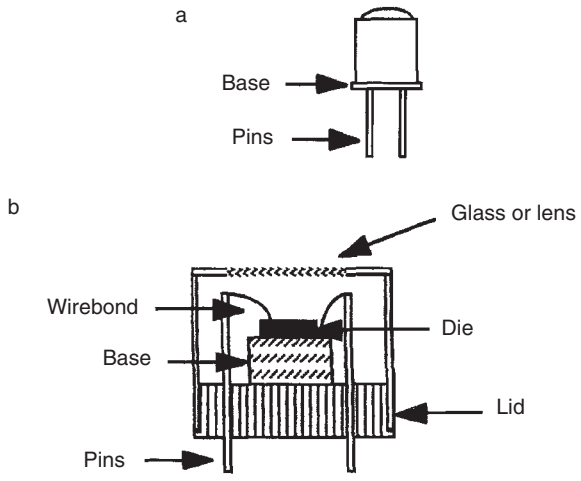


Figure 10.7 (a) TO can outline. (b) TO cross section with interconnection.

are then encapsulated using a molding process. The molded plastic then serves as the package of the device. This embodiment is very popular in the microelectronics industry because it provides for a low-cost, highly automated assembly. In optoelectronics, molded packages are less common due to the thermal and performance limitations of the devices.

The first plastic packages used for integrated circuits were introduced in the 1960s and are known as flat pacs. These packages were used to replace metal cans in which higher lead counts and surface mounting was desired. Closely following the flat pack was the introduction of the dual in-line package (DIP). Both the flat pack and the DIP are cavity packages that are wire bonded, and then the plastic is molded around the body and leads of the package. Following the introduction of the flat pac and the DIP packages, a number of other outlines were developed. Examples of flat pac and the DIP outlines are shown in Figs. 10.8a and 10.8b [12].

The key elements in a plastic package are the leadframe and the mold compound. Leadframes are typically made of copper with small amounts of iron, tin, phosphorus, or other elements to form an acceptable alloy. The design of the leadframe is optimized for manufacturing. The leadframe must accommodate the flag (which is where the die is attached), the fingers that radiate around the die (where wire bonds will be attached), and the exterior pins that are attached to the circuit board. Leadframes are typically mechanically stamped or chemically etched from copper sheets (though other alloys such as alloy 42, a mix of nickel

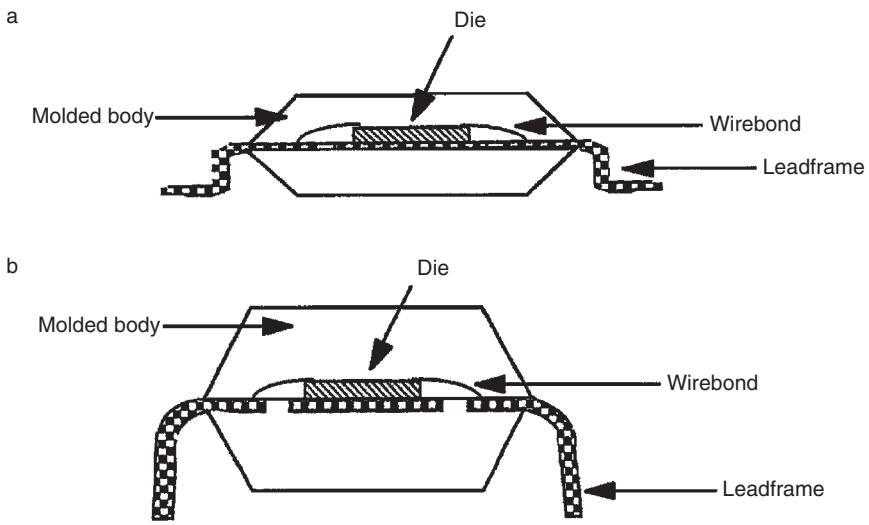


Figure 10.8 (a) Cross section of a flat pac. (b) Cross section of a DIP.

and iron, are also widely used). After formation the flag and bond fingers are plated with silver or gold, often utilizing a nickel barrier layer. The leadframes are formed in strips that allow for automated equipment to handle a number of packages simultaneously.

The mold materials are usually Thermoset compounds. The most common of these compounds are epoxy-Novolac-based molding materials. Compounds themselves are based on many components such as the resin, curing agent, mold release, accelerator, flame retarder, filler, and colorant. The mold compound must adhere well to both the leadframe and the surface of the device. Transfer molding typically preheats the resin pellet and then forces the compound into the mold cavity that holds the leadframe. The material actually flows as a viscous liquid when the viscosity drops enough due to the heat and force applied. As the resin flows through the small gates of the mold cavity it is heated through friction. This heating allows the material to cure and compact in a relatively short time. Curing of the plastic compound occurs in the mold, and in some packages it also occurs in a separate process known as postmold cure.

Some of the key concerns with molding around a device include how the mold compound affects the wire bonds that are present. Wire sweep is one of the major causes of defects in molded plastic packages. While the mold compound is filling the cavity, the flow front of the material can displace the wires so that shorting or even wire fracture can occur. Careful design of the leadframe, cavity, and wire bond placement can minimize this effect. Other concerns with plastic package manufacture center around residual stresses of the molded body, voids in the mold compound, mold/device interface or mold/flag interface, and moisture-related issues. The susceptibility of plastic packages to moisture absorption and adsorption is one of the major limiters to its applications.

10.3.3. Ceramic Packages

Ceramic packages are used extensively in the microelectronic and optoelectronic industries. They come in a variety of forms, such as DIPs, flat packs, chip carriers, and pin-grid arrays (PGAs). Many of the formats discussed in the arena of plastic packages are offered in ceramic versions and vice versa. Ceramics cover a broad area of materials that offer a variety of material characteristics. This allows the flexibility of choosing a ceramic material that can best meet one's application. Ceramics are available that have thermal expansion coefficients that match a spectra of materials, from semiconductors (3×10^{-6}/°C) to metals (17×10^{-6}/°C). Ceramics thermal conductivities range from the insulator to conductor regime (up to 220 W/m · K) and dielectric constants from 4 to 10,000 [13]. The other major advantages of ceramics are its dimensional stability and its resistance to high-temperature processes. These characteristics allow ceramics to be used for highly integrated, densely routed packages and in packages where hermetic

sealing is a requirement. In these respects, ceramic packaging services an area of applications not covered by either of the materials discussed previously.

Ceramic packages are made from a slurry of alumina, glass, and an organic binder or by utilizing dry processes and sintering of preforms. For higher density routings, ceramic packages utilize films of dielectrics and conductors on top of the alumina base layers. The two major technologies used in this area are thin film and thick film. Thin-film technology allows for higher routing density by utilizing lithographic techniques to define the conductors. Thick-film ceramic substrates are often used where separate power and ground planes are desired. Whether using thick or thin film, the ability to have multilayer packages is a significant advantage in the areas of not only routing density but also thermal and electrical distribution.

Ceramic packages are often used for "high-reliability" applications. This is based on the ability of ceramic packages to be hermetic. By hermetically sealing the lids of the package and controlling the environment of the die cavity, ceramic packages can be made to be more resistant to caustic environments than can plastic packages. One must give specific considerations to the reliability of ceramic packages. The pitfalls that arise with these package types are usually related to corrosion problems and mechanical cracking in the packages. Corrosion can be avoided by careful avoidance or removal of ionic contamination. Lid sealing of the package is also critical and must be monitored continuously in a manufacturing environment. Cracking must be managed through good manufacturing handling and careful treatment of the different materials used in the package, so as not to create a high-stress situation.

10.4. PACKAGE TO BOARD ATTACH

Attachment of the packaged device to the board is also a key consideration in the manufacturing environment. This second level of interconnection is often done to cards, boards, flexible substrates, and back planes. The key considerations at this level of attachment are its ability to be reworked, its density, and, of course, its performance. The two major types of second-level interconnection can be classified as pinned and surface mount technologies (SMT).

10.4.1. Pinned Packages

Pinned packages are a common interconnection on nearly all boards in the industry today. The most common pinned packages are the DIP and PGA formats discussed earlier. Density of a pinned package is strongly dependent on the format of the package as shown in Fig. 10.9. The density relationship between peripheral and array layouts shown in Fig. 10.9 indicates that for large pin counts area arrays allow for higher packing densities.

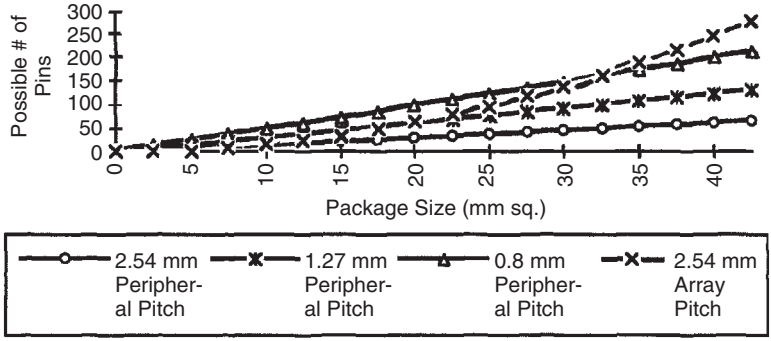


Figure 10.9 Graph of the possible numbers of pins vs. package size for a variety of pin pitches.

Pins on a ceramic package are attached using a number of different methods. The two most common methods require that the pins be inserted through the package or brazed onto the surface. Inserted pins travel through the body of the package and therefore restrict trace routing and the placement of the device to where there are no pins. In this way we have a peripheral pin layout. Brazed-on pins allow the package designer more flexibility because the device or routing metals can be directly above the pins. Therefore, brazed-on pins are often used for full-array pinned packages. Of course, in the industry through-hole pins are used in partial array configurations where a few rows of peripheral pins are used and the active routing is placed either between the pins or in areas where there are no pins present. Schematics of through-hole and brazed-on pins are illustrated in Figs. 10.10a and 10.10b.

A pinned package is attached to the board through a socket (which allows for easy replacement of the component) or, more permanently, solder attached to the board. The pins on the package are attached to the board, which has a grid of plated-through holes filled with solder. The package is then wave-soldered into place. Wave soldering allows the solder deposition and joint formation to occur in one continuous process. In wave soldering the board and the components are exposed to flux, molten solder (typically eutectic lead tin at 230–260°C) for 3–10 s and then cleaned. In this fashion multiple components can be attached to a board at one time.

One of the major advantages of the wave-solder process is that the packages are exposed to relatively low temperatures (120–160°C) for single-sided boards. The reason for this is that the exposure of the board to the wave-solder temperature occurs on the underside of the board. This limits the amount of thermal stress applied to the package.

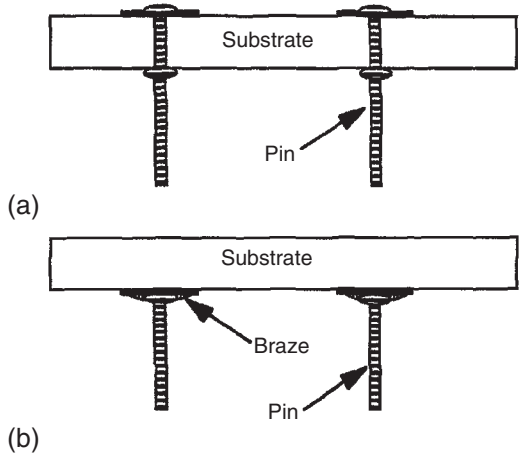


Figure 10.10 (a) Schematic of a pin through a substrate. (b) Schematic of a pin brazed onto the surface of a substrate.

10.4.2. Surface Mount Packages

Surface mount packages cover a broad number of formats as discussed previously (such as plastic lead chip carrier (PLCC), quad flat packs (QFPs), TAB, and ball grid arrays (BGAs)) and can provide a number of formats for the leads (gull-wing, J-lead, leadless, and butt joints). The use of SMT in the industry is growing due to reduced board space, cost, and the promise of higher performance solutions. Interconnection of SMT packages is considered to be more difficult and has been the focus of a great deal of industry research during the past decade. The focus of this development has been in allowing for smaller pitches (and sometimes blind mating) and in making the surface mount joints less susceptible to thermal-cycle fatigue failures. The problem of thermal-cycle fatigue is analogous to the problems discussed in Section 10.2.C.4 because these joints are typically composed of solder.

Joining an SMT to a board can be accomplished through the same wave-soldering process described in Section 10.4.A. This process is in common use for SMT packages with lead pitches above 1.27 mm but is not considered adequate for packages with finer pitches due to solder-shorting issues. The most common processes used for SMT attach are known as vapor phase reflow (VPR) and infrared reflow (IR). These processes are often preceded by an application of solder paste to the mother board and treatment of the paste prior to the introduction of the packaged components. For VPR, the board and mounted components are loaded into a chamber containing vapor and boiling liquidous perfluorocarbon compounds. Perfluorocarbons are used because they have narrow and predictable

boiling temperature ranges and are fairly inert. The assembly is rapidly brought to the desired temperature (175–250°C) and then reflowed [14].

One of the major advantages of VPR is that the processing time needed is very rapid, typically less than 5 min. This advantage is also the source of some of the significant problems with VPR. The rapid temperature rise can cause thermal shock of the components, problems with solder uniformity during reflow, and component movement due to solder wetting.

IR utilizes IR lamps providing radioactive energy to the surface components. The wavelength used by the IR lamps is usually between 1 and 5 μm. This technique is good for surface heating but causes problems with differential heating of the board surface. Differential heating is caused by different absorption of the IR energy by the components on the board. Often, the bodies of the package are heated to higher temperatures more readily than the solder lands. This problem can exasperate the temperature shock problems and lead to substantial temperature gradients across an assembled board. To overcome these shortcomings, IR is typically coupled with connective heat transfer to provide for a more stable process [15].

Other second-level interconnection technologies are also being investigated. These technologies include laser reflow, pressure interconnections, and conductive adhesives. These techniques have not found wide acceptance to date but continue to be investigated.

10.5. OPTICAL INTERCONNECT

The previous discussions incorporated manufacturing techniques that are common to both the microelectronics and optoelectronics industries. This section focuses on some of the manufacturing technologies specific to optoelectronic devices. The specific challenges encountered with packaging optoelectronic devices revolve around the transfer of optical power between the device and the outside world.

10.5.1. Coupling

In the regime of assembly techniques, we are concerned with how efficiently we can couple light into (for detectors) or out of our devices [for light-emitting diodes (LEDs) and lasers]. This coupling can be in the free space regime (e.g., CD-ROMs, bar-code scanners, or imaging) or involve the transfer of data over a fiber of any of a number of compositions. The assembly technologies used are very dependent on the application and on the level of coupling required. For coupling of devices to a single-mode glass fiber (in which the core diameter is on the order of 6–8 μm), alignment tolerances are quite tight. Coupling of devices to multimode glass fibers (in which the core diameter is on the order of 50–

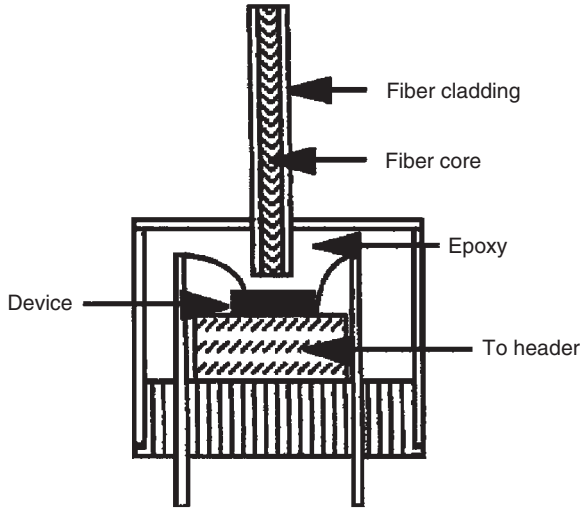


Figure 10.11 Schematic of a device that is directly coupled to a fiber.

150 μm) is less critical, and coupling to large-diameter plastic fibers (>1 mm) requires even less attention to coupling accuracy.

The manufacturing techniques that are required to meet the application's coupling needs are quite varied. The simplest method of coupling a device to a fiber is known as direct or "butt" coupling. Figure 10.11 shows a typical direct-coupled configuration [16]. This direct coupling is often used for plastic fiber and some multimode applications but is not adequate for many multimode and nearly all single-mode applications.

In addition to butt coupling, light interconnection to fibers can also be accomplished by using waveguides, lenses, and prisms. This is especially useful when manufacturing tolerances need to be enhanced, as in the case of the smaller core fibers.

10.5.2. Waveguides

Waveguides come in a number of configurations, including those made of polymers, ceramics, glass, and semiconductors. However, a waveguide can be formed from essentially any family of materials whose refractive indexes are different. In some cases, a waveguide allows for passive coupling of light to or from the device to the fiber. In other cases, waveguides are used to transfer light or to alleviate difficult tolerances.

The most commonly used waveguide is the silica fiber itself. Many of the processing techniques used for waveguide manufacture many polymer or wafer

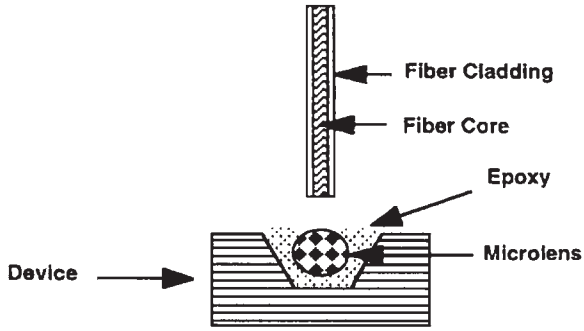


Figure 10.12 Microlens epoxied in the well of a Burrus diode.

technologies similar to those used in the manufacture of fiber and connectors. Attachment of the waveguide to the device can be directly coupled and often acts as an intermediate between the device itself and the fiber.

10.5.3. Lenses

Lenses are commonly used in the optoelectronic industry. They serve to focus and shape the light emitted from a laser or LED into a fiber. Nearly all single-mode fiber-optic products use some kind of lens to help focus and shape the beam profile for coupling to a fiber. In single-mode fiber applications, one must contend with small core diameters **(**6–8 μm) and relatively low acceptance angles (numerical aperture = 0.15, $\alpha = 8.6°$). The lenses and fiber must be placed at the focal distance from the light source to achieve the smallest spot sizes and most efficient coupling. In many cases this involves accurate placement of a semispherical lens on the surface of the optoelectronic device itself. This lens is secured through the addition of an adhesive. The fiber itself can also be secured at the focal distance from the lens by the same adhesive material. Figure 10.12 shows the use of a microlens in the well of a Burrus diode. This is often used in concert with active alignment to achieve optimal coupling. Although rigorous, this technique has provided the best solution for single-mode optical coupling in the industry.

10.5.4. Prisms

Prisms are also used in the interconnection of optical energy. In the case of prisms, one can bend light as well as selectively filter out unwanted energies. Prisms offer interesting alternatives in the selection of which radiation energies are desired. Bending light can be quite useful in coupling from fibers to devices when the end of the fiber is not in the same plane as that of the active devices.

REFERENCES

1. Anderson, O. L. 1975, November. Bell Laboratories Record.
2. Steidel, C. A. 1982. Assembly techniques and packaging. In *VLSI technology*, ed. S. M. Sze, p. 551. New York: McGraw-Hill.
3. Shukla, R., and N. Mencinger. 1985, July. A critical review of VLSI die-attachment in high reliability application. *Solid State Tech.* 67.
4. Koopman, N., T. Reiley, and P. Totta. 1989. Chip-to-package interconnections. In *Microelectronics packaging handbook*, eds. R. Tummala, and E. Rymaszewski. New York: Van Nostrand Reinhold.
5. Kawanobe, T., K. Miyamoto, and M. Hirano. 1983, May. Tape automated bonding process for high lead count LSI. *33rd Electron. Comp. Con. Proc.*, 221–226.
6. Shindo Company, LTD. Shindo TAB brochure.
7. Miller, L. E. 1969, May. Controlled collapse reflow chip joining. *IBM J. Res. Dev.* 13:239–250.
8. Fired, L. J., J. Havas, J. S. Lechaton, J. S. Logan, G. Paal, and P. Totta. 1982, May. A VLSI bipolar metallization design with three-level wiring and area array solder connections. *IBM J. Res. Dev.* 26:362–371.
9. Ohshima, M., A. Kenmotsu, and I. Ishi. 1982, November. Optimization of micro solder reflow bonding for the LSI flip chip. Second Internation Electronics Packaging Conference.
10. Greer, S. E. 1978, May. Low expansivity organic substrate for flip-chip bonding. *28th Electron. Comp. Con. Proc.*, 166–171.
11. Motorola Inc. 1983. Optoelectronics Device Data.
12. Robock, P., and L. Nguyen. 1989. Plastic packaging. In *Microelectronics packaging handbook*, eds. R. Tummala and E. Rymaszewski. New York: Van Nostrand Reinhold.
13. Tummala, R. 1989. Ceramic packaging. In *Microelectronicspackaging handbook*, R. Tummala and E. Rymaszewski (eds.). New York: Van Nostrand Reinhold.
14. Hall, W. J. 1982. Vapor phase processing considerations for surface mounted devices. *Proc. Znt. Electron. Pack. Soc. Conf.*, 41–46.
15. Dow, S. 1986, October. Infrared reflow soldering: Another approach. *Surjiuce Mount Tech.*, 28–29.
16. Midwinter, J. 1977. *Optical fibers for transmission*, 81. New York: John Wiley & Sons.

Part II

Links and Network Design

This page intentionally left blank

11

Fiber-Optic Transceivers

Michael Langenwalter
Infineon Technologies AG, Fiber Optics Division, Berlin, Germany

Richard Johnson
Infineon Technologies North America Corporation, San Jose, California

Contributions for Third Edition
Robert Atkins
IBM Corporation, Poughkeepsie, New York

11.1. INTRODUCTION

Fiber-optic transceivers (TRX), each a combination of a transmitter (Tx) and a receiver (Rx) in a single common housing, have already replaced discrete solutions in most of the datacom applications with two-fiber, bidirectional transmission in the last 12 years. The transmitter unit, with a light-emitting diode (LED) or an edge- or vertical-emitting laser diode (LD) as its radiation source, and the receiver unit, with in most cases a positive-intrinsic-negative (PIN) photodiode, are connected together to the transmission medium, either multimode or single-mode fiber. Thus, contrary to typical telecom applications with single-fiber-optic connectors at the ends of fiber pigtails, transceivers characteristically have duplex optical receptacles at one side of the housing, which fit to the corresponding duplex connectors. Depending on the application, more or less electronics with integrated circuits and passive components are implemented into the Tx and Rx units.

In this chapter we discuss the operation and application of these fiber-optic transceivers in the physical layer (PHY), along with some innovative tendencies for integration of more data transport functions into these components or for a significant reduction of size and power consumption, respectively.

The typical datacom transceiver adheres to one or more national or international standards and manufacturers may create multisource agreements (MSAs)

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

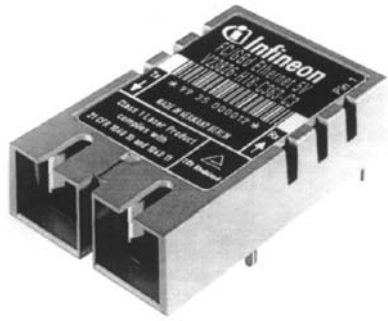


Figure 11.1 Multistandard transceiver.

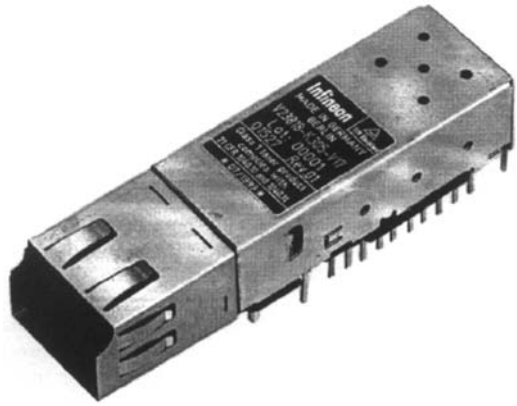


Figure 11.2 SFF GbE transceiver.

to standardize form factor, pinout, power consumption, and so on. An example standard is the ANSI Fibre Channel (FC-PI-2) for serial transceivers with data rates of 1.0625 Gbps, 2.125 Gbps, and 4.250 Gbps. The GbE standard has a data rate of 1.0 Gbps. When data rates are close, such as GbE and 1G FC, a transceiver can be designed to handle multiple standards because it is protocol agnostics (see Fig. 11.1). An example MSA is small form factor (SFF), which defines a pin-through-hole device shown in Fig. 11.2.

11.1.1. Physical Layer Interface

Transceivers are typically implemented as part of a physical link structure that incorporates data serialization. The outgoing bytes undergo 8B/10B encoding, to ensure a balanced pattern, followed by a parallel to serial conversion.

In this manner an 8-bit wide 100-MBps datastream would be converted to a 1-bite-wide 1-Gbps datastream to be delivered to the serial fiber-optic transceiver. After passing through the transmission medium, the reverse of this process is carried out.

11.1.2.1. Clock Oscillator and Regenerator

Clock generation multiplication is mainly realized using a quartz oscillator for very stable clock sources. For example, the ANSI Fibre Channel requires a clock frequency accuracy of 100 ppm. Clock regeneration/synchronization can be realized by either surface wave filter (SWF) or phase-locked loop (PLL), the latter being more common in today's applications.

11.1.2.2. Serializer and Deserializer (or Multiplexer and Demultiplexer)

The serializer/multiplexer accepts parallel data from the encoder once per byte timeframe and shifts it into the serial interface output buffers using a PLL-multiplied bit clock. The deserializer/demultiplexer accepts serial bit-by-bit data from the serial receiver output, as clocked by the recovered receiver sync clock, and shifts it back into a parallel datastream.

11.1.2.3. Encoding and Decoding

The dominant intention of coding for transmission of high-speed data is to maintain the DC balance by bounding the maximum run length of the code. Typical kinds of coding in data communication include 4B/5B coding (e.g., in FDDI) or 8B/10B coding (e.g., in ESCON/SBCON and Fibre Channel). This means high-speed receiver designs in fiber-optic transceivers are normally AC-coupled so that each DC component inside the datastream reduces the signal-to-noise ratio in the preamplifier stages.

During clock recovery, the PLLs used in today's applications require a certain edge density to ensure that the receiving PLL remains synchronized to the incoming data. In addition, word alignment can be provided by special transmission characters (e.g., a K28.5 pattern). In general, two types of characters are defined: data characters and special characters.

11.2. TECHNICAL DESCRIPTION OF FIBER-OPTIC TRANSCEIVERS

When selecting a transceiver, the key parameters are optical output power, operating data rate, receiver saturation level, receiver sensitivity, transmit and receive pulse quality (defined by rise/fall times and over/undershoot), jitter, and

extinction ratio. For transceivers, an optimal interoperation with the associated circuitry is the key to success of the application. Thus, the high-speed characteristics and immunity against external influences have to be carefully evaluated during the design phase.

11.2.1. Serial Transceivers

Serial fiber-optic transceivers are the interface between the serial electrical signal and the transmission medium, the optical fiber. They are composed of the transmitter and the receiver functions, which operate in one common housing, but are electrically independent of each other. The implementation of serial transceivers requires careful design of filter circuitry as well as sophisticated transmission line considerations for data input and output tracks to connect to the internal or additional external circuitry.

Figure 11.3 shows a simplified block diagram of a fiber-optic transceiver containing a transmitter side with a laser driver circuit and a laser diode, together with its monitor diode, for tracking the laser output. Also shown is an alarm circuitry that activates if the laser crosses a preset upper or lower optical power threshold.

The receiver side shows a PIN photodiode as optoelectric converter, a transimpedance preamplifier that converts extremely low AC currents (in the range of nA) into differential voltage signals with some mV amplitude, and the buffer stage with integrated postamplifier for ECLPECL line driving.

In most of the standards mentioned in Section 11.1, such as Ethernet, FDDI, ATM, and Fibre Channel (FC) including ESCON/SBCON, pure serial fiber-

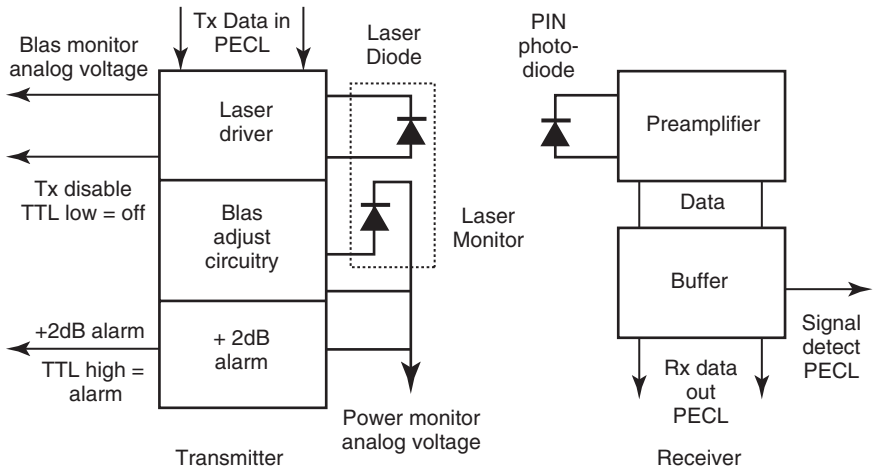


Figure 11.3 Simplified block diagram of a serial laser transceiver.

optical transceiver solutions are described, all including so-called multistandard transceivers. In order to minimize efforts in design, manufacturing, and testing of boards, a couple of parallel transceiver designs have also arisen in the interim.

11.2.2. Parallel Transceivers

Transceivers (TRX) with parallel electrical interfaces have become increasingly popular in fiber-optic link applications. Parallel TRXs provide high potential for cost saving due to

- significant board space savings by higher functional integration
- simplified board design by avoiding higher frequencies on the PC board
- low risk for development effort and design-in
- much reduced overall power consumption
- logistical benefits on all levels

11.2.2.1. Transponders

The electro-optical functions of transponders are basically very similar to parallel transceivers. But contrary to the typical design of transceivers with fiber-optic connector receptacles, transponders have connectorized fiber pigtails. They are also plugged into the board rather than soldered or surface mounted, whereas transceivers are typically directly soldered or mated to a board-mounted connector.

Figure 11.4 shows as an example the OC-48 transponder, which supports ATM/SDH/SONET applications by converting four parallel LVDS lines of

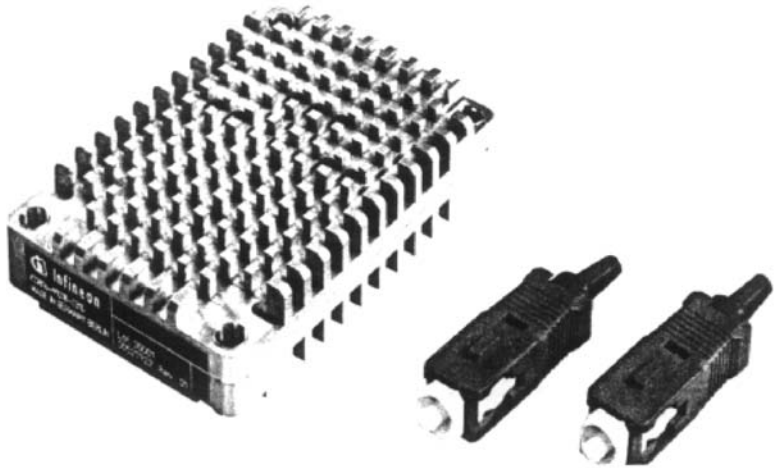


Figure 11.4 The 4:1 OC-48 transponder.

622.08 Mbit/s input/output datastream (equal to the STM-4/OC-12 hierarchy) into 2.48832-Gbit/s serial optical output/input data stream (equal to the STM-16/OC-48 hierarchy) [10–12]. This transponder is also compliant with Proposal 99.102 of the optical internetworking forum OIF.

Due to high functional integration with serializing/deserializing and, in addition, 8B/10B coding/decoding, the power consumption is the order of 3 W. Thus, heat dissipation is a serious challenge that can only be reliably managed by numerous cooling studs on top of the all-metal housing and by forced airflow inside the system device. A block diagram of the circuitry of the 4 : 1 OC-48 transponder is shown in Fig. 11.5.

11.3. THE OPTICAL INTERFACE

An interface is the meeting of two objects. In the case of the optical interface, the fiber comes together with the optical transceiver elements: the transmitter diode and the receiver diode receptacle unit. The operating link demands that the mechanical and optical properties of these two objects match. These attributes are defined in various standards to maintain a reliable interoperability of products of different manufacturers [5–8].

11.3.1. Fiber-Optic Connectors and Active Device Receptacles

A brief, closer look shows that a transceiver has not one but two optical interfaces: first, the connection of the outgoing fiber at the output of the transmitter, and second, the connection of the incoming fiber at the input of the receiver. The optical subassembly of the transmitter launches the radiation into the fiber, which will be taken up by the optical subassembly of the receiver at the other side of the link.

For a flexible configuration or a reconfigurable network, it is necessary to have repeatedly connectable and disconnectable devices. This function is realized by a single, duplex, or multifiber connector at the end of the fiber-optical cable and a port at the transceiver that accepts the fiber connector. There are various different design families on the market, such as the EC, ESCON/SBCON, FC/PC, FDDI MIC, LSA (screwed), LSG (push-pull version of LSA), LSH (E2000), SC, SMA, ST/BFOC, LC, MF, MLT (mini-SC), MT, MT-RJ, and VF-45 international standardized connector families[8]. Most of the single-fiber connectors suitable for single-mode fiber application are available in either a version with physical contact (PC) or angled polished contact (APC). Nearly all push-pull or latch-type single-fiber connectors are also available in a duplex version. ESCON/SBCON and FDDI MIC were originally duplex connectors.

Most connector systems use a ferrule-a cylindrical tube containing the fiber end that fits within a precise mechanical tolerance into the transceiver's optical

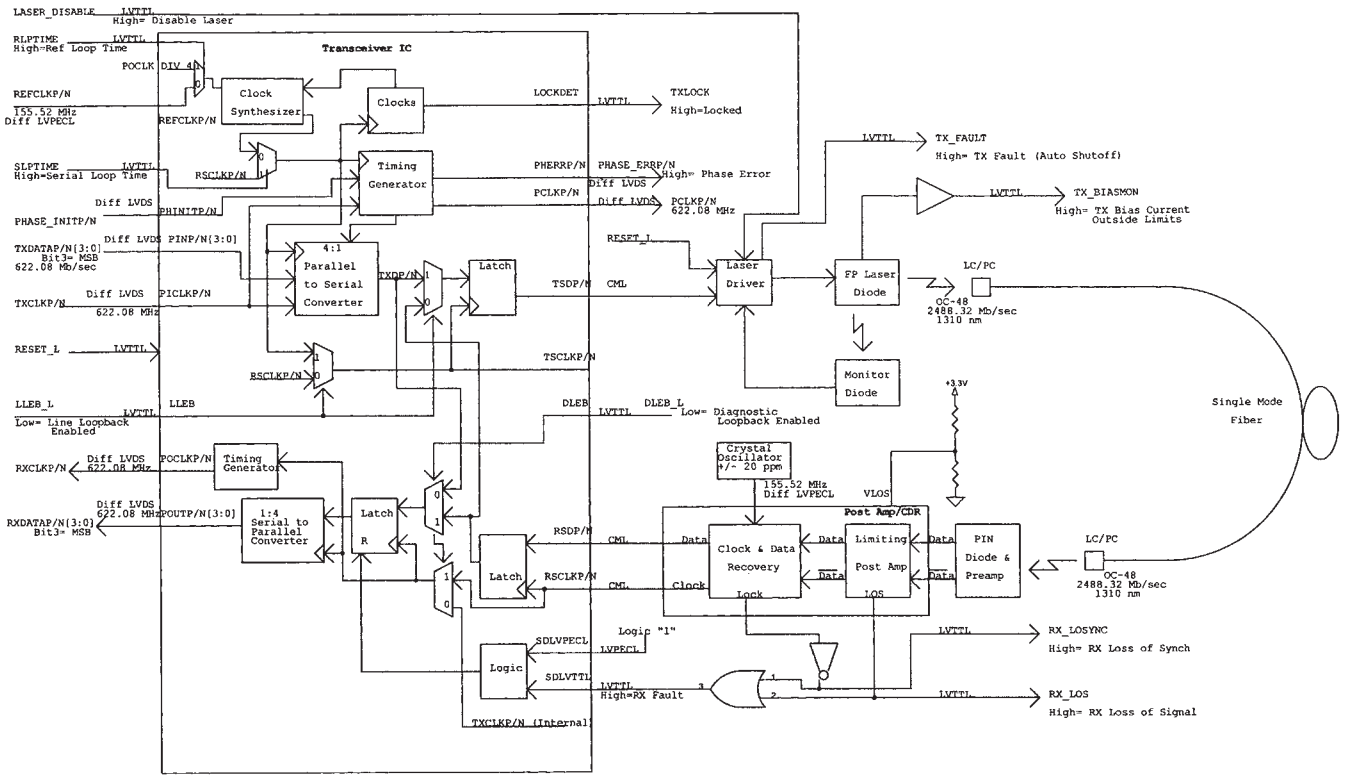


Figure 11.5 Block diagram of the 4:1 OC-48 transponder.

receptacle. The VF-45 connector is the only one with a totally different concept based on V-grooves instead of ferrules.

For ferrule-based systems, the most popular diameter of the cylindrical ferrule is 2.5 mm, while 1.25 mm has established itself for small form factor (SFF) designs such as MU- and LC-type connector families. The SMA connector is the only one with a 3.125-mm ferrule diameter. However, a ferrule diameter tolerance within a few microns must be maintained in all cases in order to achieve a satisfactory coupling of light and to ensure intermateability to receptacles of different suppliers.

Besides these single and duplex connectors, it is also possible to integrate 12 fibers into one connector to be used for parallel optical links (see Section 11.8). The most popular multifiber connectors are based on the MT design with its rectangular-shaped plastic ferrule. In general, any fiber connector has to be locked to the receptacle, for example, by screwing on a coupling ring (e.g., LSA) or snapping on a bayonet (ST) or a latch mechanism. Moreover, some connectors can be provided with a key to distinguish multimode and single-mode connectors.

For bidirectional links in datacom applications, cross-plugging must be avoided by using asymmetric-shaped connector or corresponding key structures both inside the receptacle and at the connector. In the case of parallel links using MPO connectors for a 12-lane transmitter and separate 12-lane receiver, the two 12x cables are often bundled together and the ends are labeled TX and RX.

Pigtail solutions are preferred for telecom applications. These modules do not have an optical port, but instead have one or more permanently fixed fibers of short length (i.e., on the order of 1 meter). In this case, the optical interface of the transceiver/transponder is situated at the connectorized end of the pigtail (see also Section 11.2.2.1. and Fig. 11.4).

11.3.2. The Optical Fiber

When considering the optical aspects of the interface, it is useful to look at the parameters of the fiber, the transmission medium between transmitter and receiver. An optical fiber can be described by the properties of core diameter, refractive index profile and numerical aperture, bandwidth distance product, and attenuation. These interdependent parameters more or less directly influence the requirements for the output of the transmitter and the input of the receiver.

The core diameter of the fiber defines the maximum waist of the output beam at the transmitter, because only light inside this waist will be guided by the fiber. This is why the numerical aperture, which is determined by the index profile of the fiber, limits the maximum divergence of the input beam. Launched radiation exceeding this maximum divergence will be lost. An efficient optical coupling

between transmitter and fiber can only be achieved if the core diameter matches the beam waist and the numerical aperture matches the divergence. At the receiver side, the optical coupling is mainly influenced by the waist and divergence of the beam leaving the fiber. The optical subassembly of the receiver must focus this beam onto the sensitive area of the photodiode (see also Chapter 5).

In some cases, it is necessary to consider not only the coupling efficiencies at the optical interfaces of the transceiver but also the fraction of light that is reflected back to the light source. Lasers can be very sensitive to backreflections that can disturb the laser emission and cause some noise on the optical signal. To prevent these effects, the return loss at the optical interfaces has to be high. The backreflections can be kept low by using angled polished physical contacts at the connectors or by inserting an optical isolator in front of the laser. Another method uses a design wherein the axis of the laser beam is slightly tilted to the axis of the optical fiber.

The fiber bandwidth is determined by the modal dispersion and the chromatic dispersion. Modal dispersion occurs in multimode fibers and causes a pulse broadening due to the different propagation speeds of the different paths (modes) taken by light traveling down the fiber. The excitation of fiber modes depends strongly on the optical coupling at the transceiver. Due to its index profile, single-mode fiber only supports the fundamental mode of propagation and therefore does not suffer from modal dispersion.

Pulse broadening caused by chromatic dispersion requires control of the spectral bandwidth of the transmitter. For low bit rates and short distances, LEDs with a typical full width at half maximum (FWHM) of 40 to 60 nm are sufficient. However, higher bit rates and longer distances require the application of lasers with a typical spectral width of a few nanometers down to the subnanometer region, which is achieved by distributed-feedback (DFB) lasers with only one spectral mode.

The transparency of the fiber material at the different wavelengths of radiation determines the spectral parameters of transmitter and receiver. In the past, three "classical" optical windows at 850 nm (first window), 1310 nm (second), and 1550 nm (third) of fused silica fibers were mostly used with only about ±30-nm wavelength deviation from the window's center (Fig. 11.6).

Today, these windows are broader due to the reduced OH concentration in fused silica fibers. With the availability of new types of emitters—dense wavelength division multiplexers/demultiplexers (DWDM) and different optical amplifiers (0A)—the new wide window from 1440 to 1625 nm is now applicable especially for high-speed long-haul telecom systems. The wavelength region between the second window and the fifth window, named 2e-window, may also be used in the near future by the application of special "zero-OH" fibers. Moreover, other materials—for example, polymethylmethacrylate (PMMA) plastic optical fibers (POF)—open new windows in the visible region.

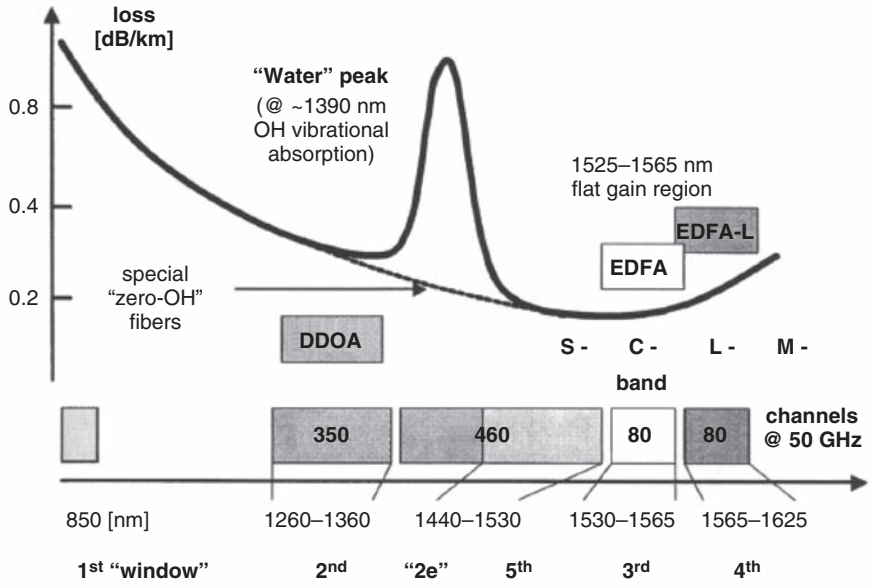


Figure 11.6 Typical spectral attenuation of silica optical fibers.

Basically, the wavelength of the transmitter and the spectral sensitivity of the receiver have to fit with these windows of low attenuation. The attenuation of the fiber (as well as any intermediate connectors), together with the sensitivity of the receiver, determines the minimum launched power in the fiber. This defines the minimum optical power of the transmitter at the optical interface if the coupling efficiency of the transmitter is provided.

11.4. NOISE TESTING OF TRANSCEIVERS

In everyday life, the dramatic increase in the use of electronic devices in general and mobile telecommunication components in particular has evoked a strong demand for electromagnetic compatibility (EMC) and electromagnetic immunity (EMI) of electronic devices. With growing efforts in the field of high-speed data transmission along with ever-increasing computer or processor integration and speed, this demand has also gained major influence on design, production, and testing of today's fiber-optic components and transmission equipment.

In close cooperation with computer, telecom, and datacom equipment manufacturers, active fiber-optic components of high EMCEMI performance have been designed. As an important foundation for success, a common understanding of

the really relevant phenomena had to be reached, and reliable and repeatable test methods had to be evaluated. This gave the basis necessary to define, then to test and to meet standardized EMCEMI requirements for system application of fiber-optic transceivers, and finally to guarantee reliable operation of these components.

This section describes established measurement procedures for the EMC/EMI performance of active fiber-optic components. The methods presented here deal with noise on supply voltage, with emission of electromagnetic noise into the surrounding environment, and with immunity against electromagnetic noise from outside.

11.4.1. Description of the Device Under Test

For all described measurements, the device under test (DUT) is a fiber-optic transmission module, for example, a transceiver or a combination of a transmitter and a receiver. The DUT is connected to a printed circuit board (PCB), which is populated with power supply filtering elements. The electrical data output lines of the receiver are connected directly to the electrical data input lines of the nearby transmitter. The top and bottom layers of the PCB are connected to ground. Connections to the measurement setup are fiber-optic jumper cables and power supply lines.

11.4.2. Noise on Vcc

Fiber-optic transceivers usually operate in an environment with fast-switching circuitry. This results in high-frequency noise on the DC power supply Vcc. High-frequency noise on Vcc cannot always be sufficiently suppressed by capacitors or inductors because these components show resonant behavior of their own, with resonant frequencies lying close to the noise frequencies. Therefore, it is essential for fiber-optic transceivers to be able to withstand high-frequency noise on Vcc. Another noise effect, commonly known as ripple, is less harmful to fiber-optic modules because ripple frequencies are lower than frequencies of noise caused by fast switching.

To perform a test on a DUT's behavior concerning noise on Vcc, a noise signal has to be coupled into the Vcc line using a bias tee. When transceivers are tested, the receiver part should act directly as a driver for the transmitter part.

If the DUT contains internal decoupling elements, noise of certain frequencies may be sufficiently suppressed. In these cases, noise voltages at Vcc device pins are close to zero. At frequencies at which noise suppression is insufficient, noise voltages can be measured at the pins.

At frequencies that are sufficiently suppressed, high noise generator amplitudes would be necessary to keep the noise voltage at the pins at a constant level. Testing a DUT under such rather artificial conditions is overly demanding. It is

therefore recommended to merely keep the noise generator amplitude at a constant level for all frequencies.

For the measurement, regular data transmission is established and then optical power is attenuated to a level that causes a certain bit error rate (BER) value. When noise is applied on Vcc, the BER increases. This increase in BER is a measure of the DUT's sensitivity to noise. The optical power is then increased until the DUT reaches the former BER. The device meets the requirements if the difference in optical power does not exceed a defined level. The measurement is repeated for different noise frequencies. It is always important to use identical test setups to ensure that measurements are consistent.

11.4.3. Electromagnetic Compatibility

11.4.3.1. Emission

The emission of electromagnetic radiation may cause improper operation of components located near a radiating device. Therefore, maximum allowable electromagnetic noise levels (depending on noise frequency) are defined. For information technology equipment the emission limits are defined in standard EN 55022, derived from IEC CISPR 22, and in FCC CFR 47, part 15, classes A and B.

The measurements are done in an anechoic chamber up to 40 GHz for any defined orientation of the DUT, or for continuous 360-degree rotation of the DUT on a wooden turntable. During testing the DUT is battery powered, and the data pattern used is a Fibre Channel test pattern at nominal data rate. The DUT's emission is measured in a range of 250 MHz up to 18 GHz and, if necessary, up to 40 GHz. The DUT's emissions are tolerated if noise remains 6 dB below these limits in the appropriate range.

11.4.3.2. Immunity

When external electromagnetic fields are applied to the DUT, bit errors can occur. The most sensitive part of a DUT is the receiver section around the transimpedance amplifier. In this section, very-low-input currents (on the order of nA) are converted into low-voltage signals. In typical transceiver designs, this part of the receiver circuitry is located near the fiber-optic connector receptacle.

Two measurement methods are established to determine the DUT's immunity: radio frequency (RF) immunity and electrostatic discharge (ESD) immunity.

11.4.3.2.1. Noise Immunity Against Radio Frequency Electromagnetic Fields

This immunity test can be performed based on IEC 61000-4-3 standard. At the start of the measurement, no noise is applied. Regular data transmission is

established (27-1 pseudorandom bit sequence—PRBS—at nominal data rate), and then the optical attenuator is set to a value at which a bit error rate of 10-6 is detected. At this operating point the DUT is more sensitive to external noise than during normal operation. The sine wave noise carrier source is switched on and regulated to generate a field strength of 3 or 10 V/m root mean square (RMS) at the DUT's location. Then amplitude modulation is applied to the carrier signal (modulation factor 80%, sine wave modulation 1 kHz). The noise carrier frequency starts at 80 MHz and is increased in steps of 1 or 10 MHz up to 2 GHz. The DUT is tested at its worst case position (evaluated experimentally) regarding its orientation and polarization of the noise field. Bit error rate (BER) values are measured for the different noise carrier frequencies.

Test results show that BER values depend on the noise frequency applied. Certain values of BER can be achieved by reducing optical power, so there is a corresponding optical power level for each BER value. Using this correspondence, test results can finally be presented as the minimum optical power necessary to achieve a certain constant BER value depending on the noise frequency. The DUT meets the requirements if the sensitivity values given in the DUT's specification are reached for all noise frequencies.

For a more detailed investigation, the difference between the minimum optical input power in the undisturbed case and the worst measurement result with noise applied can be required to be lower than a specified value. This measurement can be used in the development process to indicate improvement or degradation of a certain design.

11.4.3.2.2. Immunity Against Electrostatic Discharge (ESD)

As with most electronic devices, transceivers are sensitive to ESD [14]. ESD immunity tests deal with noise caused by fast transients from electrostatic discharges. A transient electromagnetic field based on the human body model is generated and influences the module behavior with respect to bit errors. Two tests are proposed: The first one is based on EN 61000-4-2 = IEC 61000-4-2, and the second one is suitable for a more detailed analysis. In both cases, the source of the discharges is a so-called ESD gun.

The DUT is operated with a 27-1 PRBS at nominal data rate. The optical input power is set to the lower specification limit. The DUT's receptacle is fed through a grounded plate simulating a rack panel in a possible application environment. The following are the two steps involved in this test:

1. *Without vertical coupling plate*: Electrostatic discharges (both polarities) at any point on the rack panel or at the receptacle of the DUT are applied to induce bit errors during a defined time interval of discharge. After the discharge, the DUT continues error-free operation.

2. *With vertical coupling plate*: Electrostatic discharges are applied at any point on the vertical coupling plate and thereafter to the horizontal coupling plate to induce bit errors.

The discharge voltage and the number of bit errors in the time interval observed describe the ESD immunity of the DUT.

11.5. PACKAGING OF TRANSCEIVERS (TRX)

11.5.1. Basic Considerations

As also discussed in Chapter 5, the most serious task of any packaging is to maintain the functional integrity of a device over its total lifetime under all specified conditions. One of these tasks is to protect the electronic circuitry inside the transceiver from electromagnetic influences radiated or conducted into the module as well as to prevent electronics outside the module from being influenced by the module itself. This EMI shielding may be performed by means of a conductive metallic housing, or a conductive plating on a nonconductive housing, or even separate shielding structures inside the housing. Naturally, the electronic circuitry itself should also be carefully decoupled by properly selected capacitors and other filter components where needed (see also Section 11.4.2).

The next task is to avoid possible damage from corroding chemical agents and/or humidity to sensitive devices and surfaces during the card assembly process or during normal operation. Otherwise, for example, corrosion effects in the presence of ionic impurities and humidity will occur. These effects are particularly pronounced with electronic circuitry where electromigration in the presence of DC power occurs. Therefore, the appearance of intermittent short circuits or even other catastrophic failures due to complete destruction of joints or components is almost inevitable.

Another task of the TRX housing is to avoid the influences of external mating or withdrawal forces applied by the fiber-optic connector or the cable. This is a question of mechanical stability based on design, materials, and techniques used for assembling the transceiver itself and the transceiver to a motherboard. Some typical features may be explained in more detail by the following examples.

Figure 11.7 shows the second design generation of the ESCON/SBCON transceiver, with a housing containing a transmitter optical subassembly (OSA) and a receiver OSA. Also included is a chip-on-board (COB) electronic circuit with a driver IC for the infrared radiation emitting diode (IRED) on the transmitter side and a photodiode preamplifier IC and a separated buffer IC on the receiver side. The plastic housing has ground pins that are directly molded in during injection molding. After molding, the housing is plated with copper and nickel and a flash of gold, including the ground pins. Therefore, in combination with the metallized

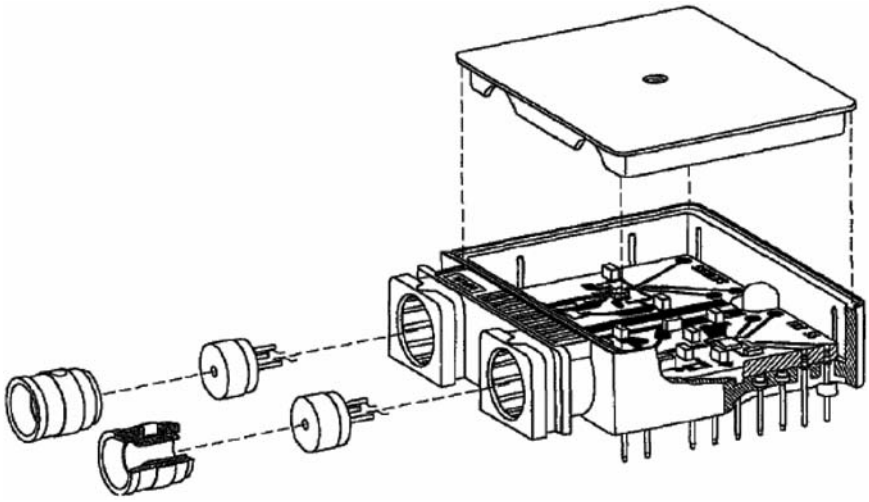


Figure 11.7 Exploded view of serial ESCON transceiver (connector shell not shown).

plastic cover, a completely closed metallic shielding and an extremely low ohmic-resistance connection to ground pins is achieved. Also, the Au-flashed nickel avoids oxidation or corrosion effects on any surface of the housing and the pins, which guarantees proper solderability.

Functional pins are press-fitted into the printed circuit board (PCB) and penetrate corresponding holes in the module housing's bottom. Also, the OSAs are press-fitted into noses on the front side of the housing. The PCB and cover are attached to the housing by conductive epoxy. To achieve washable seal, the housing's underside, as well as the nose area, are sealed by a suitable epoxy potting material.

11.5.2. Techniques for Assembling Transceivers to Motherboard

The transceiver shown in Fig. 11.7 is a typical example for a module under assembly by classic through-hole technique to a motherboard. After insertion of the module pins into the corresponding holes in the motherboard, the connector shell or the transceiver housing is screwed to the board. The transceiver is then connected to the board by a common wave soldering or fountain soldering technique followed by water-jet washing, rinsing, and hot-air blower drying. During this procedure the optical ports in the connector receptacle are protected from processing chemicals and water by a process plug.

The surface mount technique (SMT) is especially used for transceivers with higher pin counts. Also, the possibility of having components on both sides of a motherboard is a basic advantage of this technique.

Gull-wing leads in principle allow partial soldering techniques such as hot ram (thermode) or hot gas or even laser soldering. The important advantage of these processes is that only the leads are heated up to the liquid temperature of the solder and not the transceiver itself. Also, no washing procedure after soldering is needed if less aggressive fluxes are used. Moreover, a visual inspection of the solder joint quality is easily possible.

The main reason J-leaded transceivers are not widely established is their basic disadvantage with respect to partial soldering techniques and inspection. This is also a principal concern for the ball grid array (BGA) technique for transceivers with high external lead counts due to very high functional integration.

In recent years, however, the market has clearly tended toward pluggable connection of transceivers to motherboards or backplanes. An obvious advantage is the possibility to change a failed transceiver quickly and easily without any special tools. The transponder mentioned in Section 11.2.2.1 is an example of such a solution. Other examples are described in Sections 11.5.4 and 11.5.5.

11.5.3. Washable Sealed Module Design

The previously mentioned soldering processes with washing and rinsing suggest a module that should be designed for hermeticity or a washable seal, which protects against intrusion of aqueous liquids during washing/rinsing. That also means that, on the one hand, no additional care must be taken to protect single electronic devices inside the module, which certainly will reduce running cost. On the other hand, the total design and nearly all assembling agents and processes have to be adapted to the demands of the soldering process. For example, the application of a standard soldering process with aggressive fluxes may be strictly forbidden inside a sealed package. In addition, during manufacturing, a suitable and cost-saving procedure for testing the seal must be implemented in a production line. The transceiver shown in Fig. 11.7 is an example of a washable sealed design.

11.5.4. Open Module Design

The alternative to the sealed module is an open design. Such designs are well established worldwide and are becoming increasingly popular. This design is based on the simple consideration that any possibly damaging gaseous, liquid, or dusty agent going into an open housing may also be washed out or evaporate. This assumption may in general be correct, but there are design complications and challenges to open packages.

In open packages, individual protection of single electronic devices such as ICs by separate housing or encapsulation by, for example, globe top is necessary in the case of a chip-on-board (COB) technique. In addition, special care must be taken regarding the circuit layout and the circuit assembly and protection in order to prevent electromigration effects. Finally, the EM1 shielding must be managed individually by separate structures such as metallic sheets over sensitive areas of the circuitry.

With open module design, the transceiver housing itself may now be a simple injection-molded plastic part, and the final assembly of the TRX could be done by a snap-together technique. No additional sealing or corresponding test procedures are necessary.

11.5.5. Open Card Module Design

As mentioned previously, open card module designs are taking their place in fiber-optic datacom. The mechanical basis of a card module typically is a PCB combined with a peripheral plastic frame. This frame also contains the optical subassemblies (OSAs) (see also Chapter 5) and the connector receptacle on its front side. For the design of such open cards, the considerations for the previously mentioned open module designs generally apply.

Figure 11.8 shows as an example the Gigalink Card (GLC), which is a laser-based transceiver card compliant to ANSI FC-0 100-SM-LL-I standard [5] with 1.0635 GBd at 1300 nm for single-mode fiber application and 20-bit parallel

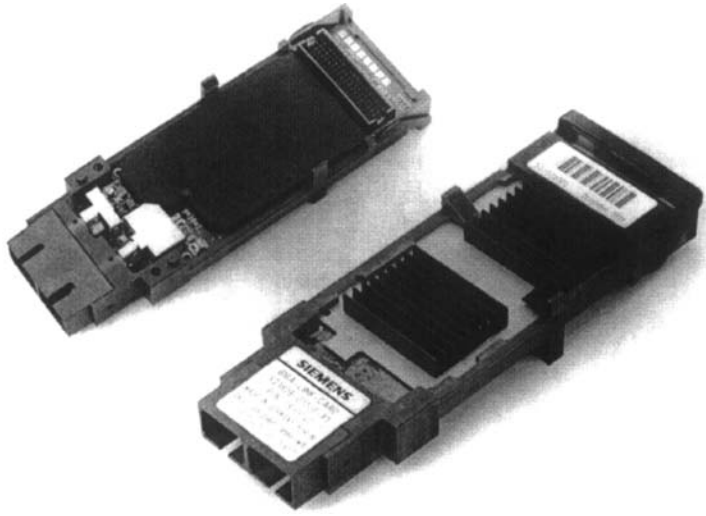


Figure 11.8 The Gigalink Card (left: bottom view; right: top view).

electrical TTL idout. Therefore, the GLC works as a transponder. The GLC is connected to a motherboard by an 80-pin connector on the rear area of the card's bottom side and secured by a pair of snapping legs. Total insertion and withdrawal forces for that connector are on the order of nearly 100 N. Therefore, a special withdrawal lever is designed as an integral part of the frame in order to prevent damage of the module card and the motherboard during withdrawal apply.

The GLC developed during the first half of the 1990s for storage area network (SAN) applications in mainframe computer systems represents the first product worldwide with the described properties.

Another pluggable module card design family is the Gigabit Interface Card (GBIC), which is becoming increasingly popular. Characteristic for the GBIC design is that the insertion/withdrawal direction is identical for both the optical and the electrical connection.

11.5.6. Small Form Factor Pluggable Design

The clear tendency toward pluggable transceivers mentioned in Section 11.5.2, combined with the other clear tendency toward increased port densities mentioned in Section 11.1.1.1, leads to the small form factor pluggable (SFP) design family. The SFP transceivers have at the bottom of the rear end an electrical connector specially designed to support hot pluggability. This means that during insertion in the motherboard, first the signal ground, then the power lines, and finally the data lines are connected. Disconnecting works in inverse order. Therefore, replacement of failed devices is possible without withdrawal of the motherboard or shutdown of the datastream of other links in operation on this motherboard.

11.6. SERIES PRODUCTION OF TRANSCEIVERS

11.6.1. Basic Considerations for Production Processes and Their Reliability

For every series production, a basic precondition for cost savings and high-quality output is a continuously running production with qualified equipment as well as qualified and reliable processes. Correct equipment and processes mean that within a defined variation of process parameters, the so-called process window, a uniform high-quality result is achieved with a high yield. Naturally, a process is safer as its window is made wider. Therefore, all equipment and processes must be qualified before the start of series production. This will typically be performed by careful and adequate testing of a statistically significant number of parts (i.e., 40 pieces at the absolute minimum) that had previously been produced on the related equipment by the corresponding process. The test results must be evaluated in a statistical manner in which the number of tested parts defines the confidence level of the result.

11.6.1.1. Qualification of Processes, the Process Capability

A brief example is provided as an explanation: A mechanical part A shall be attached to another part B by an epoxy adhesive. The key parameter for the process may be the mechanical strength characterized by, for example, the pull force needed for destruction of the epoxy attachment between parts A and B. During tests the values for the forces are monitored, and then their mean value, F_{mean}, and the standard deviation, σ, are calculated. These values are now put into a simple equation for calculating the so-called process capability:

$$C_{pk} = \frac{F_{mean} - F_{min}}{n\sigma} \geq S$$

where S is the capability index, F_{min} is the minimum force specified for the attachment, and $1 \leq n \leq 6$ is the value for the statistical sharpness. For $n = 3$, for instance, S is 1.33 in order to reach a statistical safety of $\pm 3\sigma$ or 99.73%, or a failure probability of 2700 ppm. In the case of $n = 4$ and $C_{pk} \geq S = 1.66$, the statistical safety will be $\pm 4\sigma$ or 99.9937%, or the failure probability will be 63 ppm. Therefore, an improvement of a process is simply determined by increasing its C_{pk}.

The statistical sharpness indicated by n and the correlated value of S are defined indirectly by the requested product reliability or the permitted maximum number of failures in time (FIT) in the product specification, respectively.

11.6.1.2. Correlated Environmental Tests on Subcomponents

Here it makes sense to combine the evaluation of process capability with environmental tests that are also typically defined in a product specification. These tests normally assess long-term temperature and/or humidity stress, temperature cycling, shock stress, mechanical shock, vibration stresses, and also some additional stresses determined by the customer's special application. Upper and lower stress limits or values are mostly defined by commonly known MIL or Telcordia (formerly Bellcore) or IEC generic standards named in the product specification.

The major advantage in performing tests as early as possible during the development phase and on the subcomponent level is clear: Any weakness of a construction detail, a subcomponent, or a manufacturing process will be detected earlier, resulting in time and cost savings, and improvement or general change may be more easily implemented in the design itself and in the production flow.

Naturally, all tests will also be performed at the end of product development on the completed and final product. Normally, no major weakness will occur if all pretesting has been performed carefully during the design phase.

11.6.1.3. Equipment Qualification, the Machine Capability

A similar procedure to C_{pk} may be used for calculating the so-called machine capability, C_{mk}, which is defined as the statistical safety for processes performed by a tool or a machine. For example, the uniformity of a dispensed epoxy volume will be qualified by evaluation of C_{mk} of the automated dispenser. If C_{mk} exceeds the indicated value of capability index S, then the equipment will be safe and capable of its task. A combination with environmental stress for this evaluation does not make sense.

11.6.1.4. Frozen Process

After qualification and release of equipment and a corresponding process, this production step generally should be frozen. Frozen means that in case of major changes of equipment—type or location of the equipment, process parameters, agents, or even parts—partial or total requalification of the production step must be performed. The meaning of major change or even minor change should be defined in close communication with the customer. Also, the release of a changed production step for series production typically should be with customer agreement.

11.6.2. Statistical Process Control and Random Sampling in a Running Series Production

During running production, some evident quality parameters should be fully monitored and some others checked by statistical random sampling in order to continuously assess the uniformity of processes and the quality of their output. Any deviation from the expected output quality or creeping degradation of process parameters can be detected in time for corrective action to be taken. For key process steps, a higher percentage of checks should be performed or a higher number of samples tested than would be for proven stable and uniform processes.

A commonly known tool for monitoring the quality level of any production process is the so-called statistical process control (SPC) card. Here, the actual level of a parameter is compared statistically to previously defined lower and upper limits for this most characteristic parameter that determines the quality of a component or a process. In addition, the actual level may also be compared to the so-called warning limits that typically are defined as 1σ closer than the lower and upper limit. If during production an actual value goes close to a warning limit or even exceeds it, then an early corrective action may happen immediately. Therefore, components will almost never be produced out of specification. Figure 11.9 shows an example of a SPC card.

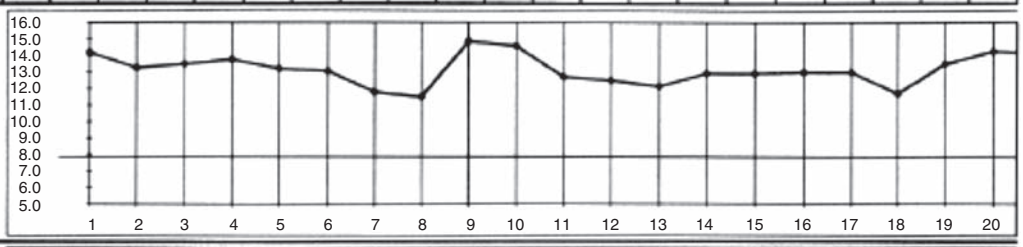
ps-No. 80	page 3.4	product	quality control card variable characteristics	part : process : parameter :

card	1	X_1																			
upper link		X_2																			
low link	5	X_3																			
unit		X_4																			
sample size	5 pcs	X_5																			
UWL		data																			
mean	11 $\bar{X}$	check																			
LWL	8	work																			
LIL	3	$\bar{X}$	14.2	13.3	13.5	13.8	13.2	13.1	11.0	11.5	14.8	14.5	12.7	12.5	12.2	12.6	12.9	13.0	13.0	11.7	13.5	14.3
mean	1.5 s	S	0.0	0.5	0.4	0.8	0.7	1.1	0.4	0.5	1.0	0.5	0.2	0.8	1.8	0.8	1.6	1.3	1.0	1.8	0.8	0.8
LIL	0	No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

test interval
start of production
change of lot of
parameters

$C_{pk} \geq 1.33$

$$C_{pk} = \frac{\bar{x} - T_1}{3 * S}$$



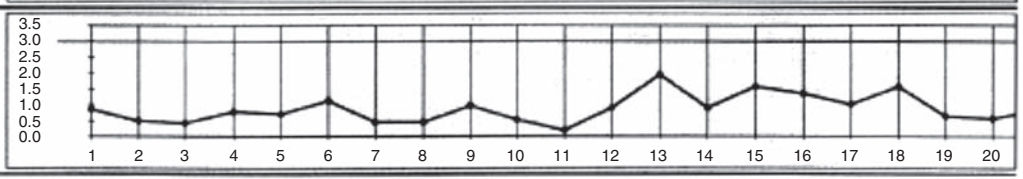


Figure 11.9 Example of an SPC card.

11.6.4. Zero Failure Quality Burn-in, Final Outgoing Inspection, and Ship to Stock

Any technical system or component behaves with respect to its failure probability according to a well-known time-dependent failure function, the so-called bathtub life curve. This curve describes the fact that a component will most probably fail at a higher rate at the early beginning of its lifetime (BOL) and very late at the end of its lifetime (EOL). In between, the probability of failure is low and constant.

A well-established method for identifying early failing parts, especially in electronics, is to perform a burn-in. During burn-in, the fiber-optic transceivers are operated at an elevated temperature level over a defined period of time. The time, temperature, and possibly some DC power overload will define the confidence level for effective screening. Before and after burn-in, the transceivers will run their normal complete inspection of all relevant electro-optical parameters, and the measured values will be compared. Individual transceivers will be rejected if there is a delta in any parameter exceeding a defined maximum value. Therefore, the failure rate for field-installed components can be reduced dramatically, and the goal of a real zero-failure quality is approached.

11.7. TRANSCEIVERS TODAY AND TOMORROW

11.7.1. Transceivers Today

Fiber-optic transceivers for applications in the field of datacom are mostly characterized by a couple of established international standards. These standards define the electro-optical performance of a transceiver/transponder as well as its pinout and its physical outline and package, including the corresponding fiber-optic connector interfaces [6, 7, 8].

Fiber-optic transceivers meeting these standards are operating worldwide in numerous applications in mainframes, server clusters, storage area networks, wide area networks, and local area networks, and currently around 20 to 30 worldwide competing suppliers have been established. The number of partners involved in some important multisourcing agreements has seen an increase since 1989. This is also indicative of the increasing importance of industrial associations where both suppliers and applicators are represented. This speeds up the market penetration of novel components, systems, and applications. Nowadays, this does not seem to generate conflicts with the commonly agreed normative power of international standardization organizations such as the International Organization for Standardization (ISO), International Electrotechnical Commission (IEC), and International Telecommunication Union (ITU).

The demand for these transceivers has continuously increased during the past 10 years, and the prices have shown dramatic decreases of the order of 25%

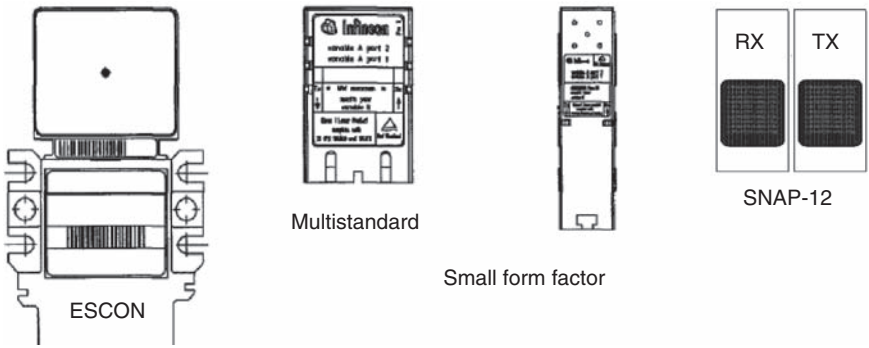


Figure 11.10 Comparison of the outlines of different transceiver generations.

per year. Consequently, the goal of all manufacturers is to offer a high level of performance, reliability, quality, and serviceability while maintaining cost-effective production in the face of drastically increased volumes to meet the market pricing.

11.7.2. Some Aspects of Tomorrow's Transceivers

The bit rates of fiber-optic transceivers are continuously increasing in order to meet the worldwide demand for ever higher bandwidths. These bandwidth increases are called for by both existing storage and networking markets, as well as the parallel computing industry and high-end server design.

11.7.2.1. Geometrical Outline of Transceivers

In the past 10 years, a significant reduction of module/transceiver size was possible due to significant progress in the downsizing of optical subassemblies (see Chapter 5) and associated passive and active electronic components and circuitry. Figure 11.10 shows an in-scale comparison of the ESCON/SBCON outline (left), multistandard, small form factor (SFF), and parallel SNAP-12 transceivers (right). The function of the transceivers shown is described in detail in Section 11.1.1.

If one combines the increase of bit rate with the reduction of size, the success of the development efforts of the past 10 years is obvious. Figure 11.11 shows a graph for the bit rate per square millimeter, named "rate-density," versus the years of introduction of the products to the market. The dots represent, from left to right: ESCON/SBCON, MS 155 Mbit/s, MS 622 Mbit/s, SFF 1 Gbit/s, and SFF 2.5 Gbit/s. The first dot differs from the last dot by a factor of 100. There is no obvious reason why this trend should change in the near future.

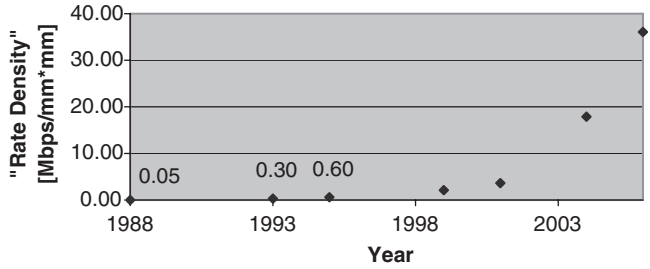


Figure 11.11 "Rate-density" of transceivers.

11.7.2.2. Functional Integration

Another direction for the next generation of transceivers is the inclusion of additional electronic functions in a common module housing, such as

- Serialization and deserialization of parallel digital bit-streams
- Encoding and decoding of serial bit streams
- Clock synchronization/regeneration on the receiver side
- Laser control and laser safety functions in laser-based transceivers/transponders as previously discussed in Section 11.2.2

The main advantage for the user of such higher levels of integration is that no high-speed signals are on the system board, with the related cost savings. One challenge developers need to solve is the issue of heat dissipation caused by such increased integration of a large number of high-speed digital electronic functions within the package. The only reasonable solution is to reduce the power consumption by application of low-power IC technologies with a supply voltage of 3.3 V or less. Such ICs have been already introduced and will be continuously improved for reduced power consumption.

An additional complication arises when some of the ICs in transceivers have to operate mixed signals, which means pure digital signals combined with analog signals, and in addition DC bias voltages and control functions. In the case of laser-driver ICs, the bandgap of the laser's active radiating material defines the absolute minimum of the supply voltage, given by fundamental physical laws. The lower the wavelength emitted by the radiation source, the higher the bandgap energy and, consequently, the higher the required bias voltage. Therefore, the supply voltages of fiber-optical transceivers may not completely follow the general tendency in digital electronics toward continuously decreasing supply voltages.

11.7.2.3. Edge-emitting Lasers and VCSELs as Optical Sources

If one exceeds the bit rate of approximately 300 Mbit/s in a fiber-optic intermediate-range multimode fiber link, the commonly used IRED on the

transmitter side will be too slow and must be replaced by a laser diode (LD). However, the very fast (up to 10 Gbit/s with direct modulation) conventional edge-emitting laser diode (EELD) is more complicated in application than an IRED because of the following:

1. An EELD needs a control circuit that monitors the output optical power and compensates for temperature and aging effects.
2. Accurate and reliable optical coupling of an EELD into a single-mode fiber is much more difficult and therefore more expensive compared to coupling an IRED into a multimode fiber. The position accuracy and stability needed for EELD in the range of 0.1 micrometer is approximately an order of magnitude higher than for coupling an IRED.
3. Laser safety due to potentially high optical output power has to be taken into account. The limitation of radiated optical power can be achieved by optical means or by electrical limitation of LD power output. Nevertheless, products for most datacom applications are unlikely to be successful in the market without certification as laser class 1 safe according to IEC 60825-1 or corresponding regulations such as those of the FDA (see also Section 11.3.3).

Currently, a new type of laser is becoming dominant in some specific applications, the vertical cavity surface-emitting laser (VCSEL). This source was originally developed as 980-nm pumping of erbium doped fiber amplifiers for long-haul telecom transmission lines.

One of the key advantages of a VCSEL compared to an EELD is the IRED-like technology. This allows one to produce VCSELs with all processing steps, including burn-in and final testing, completely at a wafer level. Some additional advantages of VCSELs with respect to the EELD are listed in Table 11.1. A disadvantage of VCSELs is that not all of the wavelength bands covered by EELDs are available with VCSEL technology. Currently, only VCSELs for the 850-nm band are available for volume production with proven reliability and lifetime. VCSELs for the 1300-nm and the 1550-nm bands are still under basic research and design development. The experts estimate that possibly in the next four years 1300-nm VCSELs will also be available in small volumes with acceptable yield.

11.7.2.4. Laser Diodes for Multimode Fibers, Mode Underfill

Worldwide there are many miles of graded-index (GI) multimode fibers installed in buildings and campuses. However, the speed and transmission field length of fiber-optic links with GI multimode fibers combined with IREDs is limited due to power budget and bandwidth-length limits.

In order to safeguard this investment and use this current cabling even for higher speed transmission over distances of more than 100 m, the concept emerges

Table 11.1
Comparison of Features: Edge-emitting laser diode (EELD) vs. vertical cavity surface-emitting laser (VCSEL).

Feature	EELD	VCSEL
Wavelength bands	650, 850, 1300 to 1660 nm	(650), 850, (1300, 1550) nm
Spectral bandwidth	Very narrow	Narrow
Size of active area	Typically 0.5–1 × 2–10 µm	Variable, 5–50 µm diameter
Beam geometry	Strong elliptic	Circular
Beam divergence	High, up to 60° × 20°	Low, ca, 5°
Number of modes	Typically 1 or few	1 or even up to many 10 s
Coupling to fiber	Difficult and sensitive	Easy
Coupling efficiency	Moderate	High
Threshold current	Approximately 10 mA	Some mA
Direct modulation bandwidth	High, up to 10 Gbit/s	High, up to 10 Gbit/s
Temperature drift of P_{opt}	Fairly high	Tendentially low
Environmental sensitivity	Extremely high	Moderate
Processing of chip	Very specific	Similar to LED
Final processing	Single bar	On wafer
Burn-in and functional test	Single on heatsink	On wafer

of using laser diodes as sources for GI multimode fibers. There are groups studying the idea of extending the limits of GI multimode fibers by means of the so-called mode underfill launch condition. That would mean that coupling of optical power from a LD with limited focal diameter and numerical aperture into a GI multimode fiber would establish only a few low-order propagation modes near the center of the fiber core. The result would be a significant increase of the fiber's bandwidth-length limit.

Experimental investigations have confirmed the theoretical assumptions. Therefore, transmission of up to 1 Gbps with 1300-nm wavelength over more than 500 m of standard graded-index multimode fibers would work well. This direction is still receiving intensive discussion in the related standardization groups.

However, this technique will establish itself only if the price and performance for laser-based products are drastically improved. One key component would be an inexpensive laser optical subassembly (see also Chapter 5) with a laser diode that operates uncooled over the temperature range of category C, controlled environment (–10°C to +60°C), according to IEC 61300-2-22, a typical office or building environment.

11.8. PARALLEL OPTICAL LINKS

11.8.1. High-Density Point-to-Point Communications

Fiber-optic transceivers have become well established for applications requiring high-bandwidth transmission of data. Such applications include backbone

switching for telecommunications, high-end routers, storage area networks (SANs), cross data center communications (Ethernet), and data flow for disk clusters.

Point-to-point communications are often configured as "patch panels," in which the fiber-optic transceivers are mounted onto a front panel, with the fiber sockets accessed through holes in the front panel. A duplex fiber cord is routed from one transceiver in one rack to the next desired transceiver in an adjacent rack. The number of fiber cords that a given panel can support, and hence the total aggregate bandwidth available from standard fiber-optic transceivers, is practically limited by the number of fiber sockets that can be installed on the panel. Consider, for example, a small form factor transceiver with a width of approximately 14 mm, operating at a data rate of 2.5 Gbit/s. Such a transceiver offers a bandwidth per front panel width of almost 200 Mbit/s per millimeter; let us call this the bandwidth density.

As bandwidth density requirements increase, the density limit imposed by single-channel transceivers becomes increasingly burdensome. This density constraint can be significantly relaxed by using a combination of multiple-channel fiber-optic modules and multifiber ribbon cable.

SANs and cross data center communication applications are stressed with more and more data every year. The high-volume serial protocols respond to this with regular increases in data rate; Ethernet has recently increased from 1 Gbps to 10 Gbps, and FICON/FC has gone from 1 Gbps/2 Gbps multirate transceivers to 1 G/2 G/4 Gbps multirate parts, with 8 Gbps coming soon. These serial transceivers allow host bus adapter (HBA) cards to be designed with several high-bandwidth ports, and directors and switches to be designed with tens of ports brickwalled on both sides of the client cards.

The parallel computing industry supplies products to meet the demands of most complex simulation and modeling problems, such as global climate modeling and protein folding. These problems are split into thousands of small chunks that are computed by individual processors. The results from, and new inputs to, the processors must be communicated through a switching fabric to keep the program moving forward, which results in very high IO bandwidth requirements from the card edge. Parallel transceivers operating are available today with single-lane bandwidths from 2 to 6 Gbps (with Double Date Rate Infiniband, DDR-IB, at 5 Gbps as one standard example) and individual transmitters and receivers housing from 4 to 12 lanes. One MSA related to such parallel transceivers is the SNAP-12 standard. Using a typical 20-mm center-to-center spacing, one can fit a transmitter/receiver pair in 40 mm of card edge with an aggregate bandwidth of 60 Gbps at DDR-IB.

11.8.2. Common Parallel Optic Module Configurations

Just as multifiber cables improve the bandwidth density of the front panel, 12-channel fiber-optic modules dramatically improve the area utilization of the

printed circuit board (bandwidth per unit board area). A transmitter module consists of a linear array of 12 lasers plus associated drive electronics; a receiver module consists of a linear array of 12 PIN diodes plus associated transimpedance amplifiers. The operation of each channel is independent of that of the next adjacent channel.

VCSELs are by far the most common choice for laser in the transmitter modules because of low cost and ease of launching laser light into the optical fiber. Currently, the state of the art is 850-nm emission of multimode light; thus the fiber cable should also be multimode. VCSEL arrays operating at 1310 nm are available from a number of manufacturers.

An alternative configuration to the 12-channel parallel optics combines four transmitters and four receivers into a single package. A 12-fiber ribbon cable is typically used, with the center four fibers "dark."

11.8.3. Link Reach

One of the most critical questions about a parallel optical link is, "What is a reasonable link reach?' This means, "What fiber cable length can be supported while still obtaining acceptable link performance in a low-cost installation?" This is a complex issue that prompts at least three distinct questions.

The first question concerns technical feasibility: What link reaches can be demonstrated in a laboratory? The current state-of-the-art is for a per-channel bandwidth of 2.5 Gbit/s at an operating wavelength of 850 nm. Such an optical signal propagating through multimode becomes degraded through one of three mechanisms: optical absorption (which is significantly higher at 850 nm than at 1310 nm), chromatic dispersion (which is much more severe at 850 nm than at 1310 nm), and modal dispersion. Optical attenuation and chromatic dispersion performances are largely defined by the glass materials system, and hence are not likely to exhibit major improvements. However, the third mechanism, modal dispersion, is highly sensitive to the fiber manufacturing process. A number of fiber manufacturers are optimizing their processes to produce very low modal dispersion fiber for operation at 850 nm. Lucent's LazrSpeed and Coming's NGMM fiber are examples of this effort. A number of companies have demonstrated excellent link performance over a reach of at least a kilometer at 2.5 Gbit/s using such fiber.

The second question concerns prudent system design: What is a reasonable link budget that gives high assurance of successful operation under essentially all circumstances? Just because a particular link reach can be (routinely) demonstrated in the laboratory does not make this a prudent choice for system design. Typically, a desired system design is expressed in terms of a link budget. The lowest expected laser power (over the life of the laser, for all allowed performance limits of temperature and voltage), minus the worst-case receiver sensitivity,

defines a possible operating range. From this range must be subtracted expected losses such as worst-case fiber connector losses and fiber attenuation. Additional "penalties" are deducted to account for degradation mechanism such as laser residual intensity noise (RIN).

Parallel optics modules are at a disadvantage compared with single-channel transceivers for achieving link budgets. Specifications on expected transmitter optical power need to be wider for parallel optics module to account for expected channel-to-channel power variations. Furthermore, laser safety limits are more restrictive for a parallel optics module because of the multiple channels. As performance goals become ever more aggressive, finding an optimum balance between laser safety constraints and a prudent link budget becomes an ever more difficult challenge for parallel optics.

The third question concerns the cost of the link, as fiber cable costs (especially for high-performance multimode ribbon fiber) can be a substantial fraction of the total link cost.

11.8.4. A Look to the Future

Cost and performance analysis changes dramatically if parallel links are configured with single-mode fiber operating at 1310-nm wavelength. Mode dispersion is eliminated because of the single-mode fiber. Optical attenuation is significantly reduced, as is chromatic dispersion. Laser safety is much less restrictive at the longer wavelength, so higher optical powers can be considered. Thus significantly longer link reach can be realized at 1310 nm. But the most dramatic change is the cable cost. Single-mode multifiber cables should have the lowest cost of all multifiber possibilities.

Such a link would require two key changes, however. One is the use of a long-wavelength laser. For reasons of cost and ease of light launch, the laser array is preferably a VCSEL. The second major change is a dramatic tightening of opto-mechanical tolerances. Single-mode fiber has a core diameter that is almost an order of magnitude smaller than standard multimode fiber. This will make the manufacture of such a parallel optics module much more challenging. While 1310-nm VCSELs have been available now for a couple of years, their adoption has been slow. Thus, single-mode 1310-nm parallel optics modules are not yet available.

Future transceiver design is likely to focus on power consumption, electromagnetic compatibility and immunity, and density. As data rates continue to increase, we will start to see transceivers used closer to the ICs on the board and not just at the card edge. It has also been demonstrated that it is possible to incorporate optical components onto a chip, completely avoiding the deficiencies of high-speed signals on copper board traces. While these advancements may take their place in high-end computing systems, classical card edge transceivers are

likely to continue to play their role into the foreseeable future to allow fiber cable connection for SANs and networking.

ACKNOWLEDGMENTS

Many thanks to all colleagues for their help in giving hints for corrections and updates, in particular:

- Thomas Murphy for careful check of grammar and wording in this chapter
- Herwig Stange for the update of currently valid laser safety limits
- Mario Festag for checking and updating Section 11.4.3
- Ursula Annbrust and Renate Lindner for their help in preparing the figures, graphs, and photos

REFERENCES

1. Agrawal, Govind P. 1997. *Fiber-optic communication systems*, 2nd ed. New York: Wiley.
2. Saleh, B. E. A., and M. C. Teich. 1991. *Fundamentals of photonics*. New York: Wiley.
3. Proceedings of 26th ECOC. September 3–7, 2000. Munich, Germany: VDE-Verlag.
4. IEC CA/l727/QP, 2000, March. SB4 FWG: Survey of future telecommunications scenario.
5. ANSI X3T9.x and T1l.x. Fibre Channel (FC) Standards incl. FDDI, SBCON and HIPPI-6400, URLs: http://web.ansi.org/default.htm and http://www. fibrechannel.com.
6. IEC SC86C Drafts, released or midterm to be released IEC Standards, Group 62 148–xx, Discrete/integrated optoelectronic semiconductor devices for fiber optic communication—Interface Standards, URL: http://www.iec.ch.
7. IEC SC86C Drafts, released or midterm to be released IEC Standards, Group 62 149–xx, Discrete/integrated optoelectronic semiconductor devices for fiber optic communication including hybrid devices—Package interface standards.
8. IEC SC86B Drafts, released or midterm to be released IEC Standards, Group 61 754–xx, Fibre Optic Connector Interfaces.
9. IEEE Projects 802.x, LAN/MAN Standards and Drafts URL: http://standards.ieee.org.
10. Telcordia Technologies (formerly BELLCORE) GR-253-CORE. 2000, September. Issue 3. Synchronous Optical Network (SONET) Transport Systems: Common Generic Criteria.
11. ITU-T G.957. 1999, June. Optical interfaces for equipment and systems relating to the synchronous digital hierarchy (SDH).
12. ITU-T G.958. 1994, November. Digital line systems based on the synchronous digital hierarchy (SDH) for use on optical fiber cables.
13. International Standard IEC 60825-1,1993 incl. Amendment 2, January 2001, ISBN 2-8318-5589-6, Safety of laser products—Part 1: Equipment classification, requirements and user's guide.
14. Atkins, R., and C. DeCusatis. 2006, March 27–28. Latent electro-static damage in vertical cavity surface emitting semiconductor laser arrays. Proc. 2006, IEEE Sarnoff Symposium, Princeton, NJ.

12

Optical Link Budgets and Design Rules

Casimer DeCusatis
IBM Corporation, Poughkeepsie, N.Y.

12.1. FIBER-OPTIC COMMUNICATION LINKS (TELECOM, DATACOM, AND ANALOG)

There are many different applications for fiber-optic communication systems, each with its own unique performance requirements. For example, analog communication systems may be subject to different types of noise and interference than digital systems, and consequently require different figures of merit to characterize their behavior. At first glance, telecommunication and data communication systems appear to have much in common, as both use digital encoding of datastreams. In fact, both types can share a common network infrastructure. Upon closer examination, however, we find important differences between them. First, datacom systems must maintain a much lower bit error rate (BER), defined as the number of transmission errors per second in the communication link (we will discuss BER in more detail in the following sections). For telecom (voice) communications, the ultimate receiver is the human ear, and voice signals have a bandwidth of only about 4 kHz. Transmission errors often manifest as excessive static noise such as encountered on a mobile phone, and most users can tolerate this level of fidelity. In contrast, the consequences of even a single bit error to a datacom system can be very serious; critical data such as medical or financial records could be corrupted, or large computer systems could be shut down. Typical telecom systems operate at a BER of about 10^{-9}, compared with about 10^{-12} to 10^{-15} for datacom systems.

Another unique requirement of datacom systems is eye safety vs. distance tradeoffs. Most telecommunications equipment is maintained in a restricted environment and is accessible only to personnel trained in the proper handling of

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

high-power optical sources. Datacom equipment is maintained in a computer center and must comply with international regulations for inherent eye safety; this limits the amount of optical power that can safely be launched into the fiber, and consequently limits the maximum distances that can be achieved without using repeaters or regenerators. For the same reason, datacom equipment must be rugged enough to withstand casual use, while telecom equipment is more often handled by specially trained service personnel. Telecom systems also tend to make more extensive use of multiplexing techniques, which are only now being introduced into the data center, and more extensive use of optical repeaters.

12.2. FIGURES OF MERIT: SNR, BER, AND MER

Several possible figures of merit may be used to characterize the performance of an optical communication system. Furthermore, different figures of merit may be more suitable for different applications, such as analog or digital transmission. In this section, we will describe some of the measurements used to characterize the performance of optical communication systems. Even if we ignore the practical considerations of laser eye safety standards, an optical transmitter is capable of launching a limited amount of optical power into a fiber. Similarly, there is a limit as to how weak a signal can be detected by the receiver in the presence of noise and interference. Thus, a fundamental consideration in optical communication systems design is the optical link power budget, or the difference between the transmitted and received optical power levels. Some power will be lost due to connections, splices, and bulk attenuation in the fiber. There may also be optical power penalties due to dispersion, modal noise, or other effects in the fiber and electronics. The optical power levels define the signal-to-noise ratio (SNR) at the receiver, which is often used to characterize the performance of analog communication systems. For digital transmission, the most common figure of merit is the bit error rate (BER), defined as the ratio of received bit errors to the total number of transmitted bits. Signal-to-noise ratio is related to the bit error rate by the Gaussian integral

$$\mathrm{BER} = \frac{1}{\sqrt{2\pi}} \int_{Q}^{\infty} e^{-\frac{Q^2}{2}} dQ \cong \frac{1}{Q\sqrt{2\pi}} e^{-\frac{Q^2}{2}} \tag{12.1}$$

where Q represents the SNR for simplicity of notation [1–4]. From Eq. (12.1), we see that a plot of BER vs. received optical power yields a straight line on a semilog scale, as illustrated in Fig. 12.1. Nominally, the slope is about 1.8 dB/decade; deviations from a straight line may indicate the presence of nonlinear or non-Gaussian noise sources. Some effects, such as fiber attenuation, are linear noise sources; they can be overcome by increasing the received optical power, as seen from Fig. 12.1, subject to constraints on maximum optical power (laser

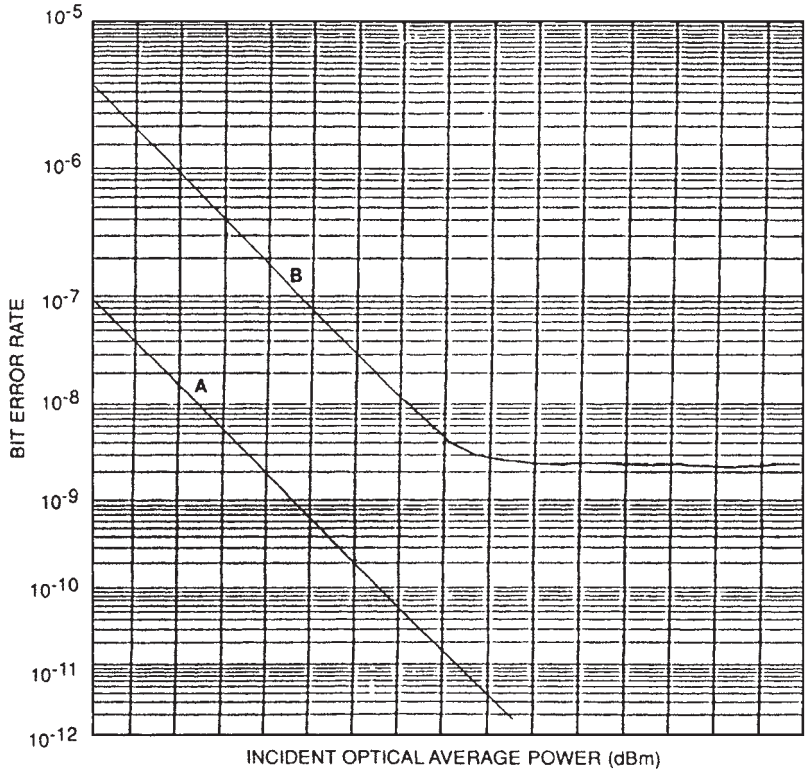


Figure 12.1 Bit error rate as a function of received optical power. Curve A shows typical performance, whereas curve B shows a BER floor.

safety) and the limits of receiver sensitivity. There are other types of noise sources, such as mode partition noise or relative intensity noise (RIN), which are independent of signal strength. When such noise is present, no amount of increase in transmitted signal strength will affect the BER; a noise floor is produced, as shown by curve B in Fig. 12.1. This type of noise can be a serious limitation on link performance. If we plot BER vs. receiver sensitivity for increasing optical power, we obtain a curve similar to Fig. 12.2, which shows that for very-high-power levels, the receiver will go into saturation. The characteristic "bathtub"-shaped curve illustrates a window of operation with both upper and lower limits on the received power. There may also be an upper limit on optical power due to eye safety considerations.

We can see from Fig. 12.1 that receiver sensitivity is specified at a given BER, which is often too low to measure directly in a reasonable amount of time (for example, a 200-Mbit/s link operating at a BER of 10^{-15} will only take one error

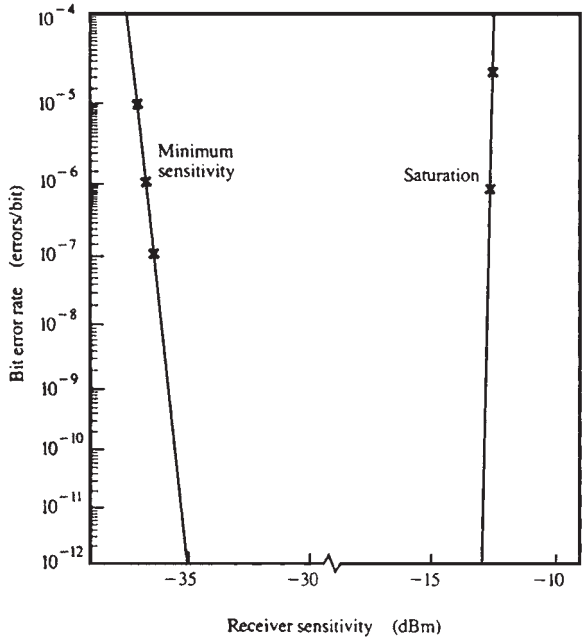


Figure 12.2 Bit error rate as a function of received optical power illustrating range *c*, operation from minimum sensitivity to saturation.

every 57 days on average, and several hundred errors are recommended for a reasonable BER measurement). For practical reasons, the BER is typically measured at much higher error rates, where the data can be collected more quickly (such as 10^{-4} to 10^{-8}) and then extrapolated to find the sensitivity at low BER. This assumes the absence of nonlinear noise floors, as cautioned previously. The relationship between optical input power, in watts, and the BER is the complementary Gaussian error function

$$\text{BER} = 1/2 \text{ erfc } (P_{out} - P_{signal}/\text{RMS noise}) \tag{12.2}$$

where the error function (erfc) is an open integral that cannot be solved directly. Several approximations have been developed for this integral, which can be developed into transformation functions that yield a linear least squares fit to the data [1]. The same curve-fitting equations can also be used to characterize the eye window performance of optical receivers. Clock position/phase vs. BER data are collected for each edge of the eye window; these data sets are then curve fitted with the above expressions to determine the clock position at the desired BER. The difference in the two resulting clock positions on either side of the window gives the clear eye opening [1–4].

In describing Figs. 12.1 and 12.2, we have also made some assumptions about the receiver circuit. Most data links are asynchronous and do not transmit a clock pulse along with the data; instead, a clock is extracted from the incoming data and used to retime the received datastream. We have made the assumption that the BER is measured with the clock at the center of the received data bit; ideally, this is when we compare the signal with a preset threshold to determine if a logical "1" or "0" was sent. When the clock is recovered from a receiver circuit such as a phase-locked loop, there is always some uncertainty about the clock position; even if it is centered on the data bit, the relative clock position may drift over time. The region of the bit interval in the time domain where the BER is acceptable is called the eyewidth; if the clock timing is swept over the data bit using a delay generator, the BER will degrade near the edges of the eye window. Eyewidth measurements are an important parameter in link design, which will be discussed further in the section on jitter and link budget modeling.

In the design of some analog optical communication systems, as well as some digital television systems (for example, those based on 64-bit quadrature amplitude modulation), another possible figure of merit is the modulation error ratio (MER). To understand this metric, we will consider the standard definition of the Digital Video Broadcasters (DVB) Measurements Group [5]. First, the video receiver captures a time record of N received signal coordinate pairs, representing the position of information on a two-dimensional screen. The ideal position coordinates are given by the vector (X_j, Y_j). For each received symbol, a decision is made as to which symbol was transmitted, and an error vector $(\Delta X_j, \Delta Y_j)$ is defined as the distance from the ideal position to the actual position of the received symbol. The MER is then defined as the sum of the squares of the magnitudes of the ideal symbol vector divided by the sum of the squares of the magnitudes of the symbol error vectors:

$$\mathrm{MER} = 10\log\frac{\sum_{j=1}^{N}(X_j^2 + Y_j^2)}{\sum_{j=1}^{N}(\Delta X_j^2 + \Delta Y_j^2)}\mathrm{dB} \tag{12.3}$$

When the signal vectors are corrupted by noise, they can be treated as random variables. The denominator in Eq. (12.3) becomes an estimate of the average power of the error vector (in other words, its second moment) and contains all signal degradation due to noise, reflections, transmitter quadrature errors, and so on. If the only significant source of signal degradation is additive white Gaussian noise, then MER and SNR are equivalent. For communication systems that contain other noise sources, MER offers some advantages; in particular, for some digital transmission systems there may be a very sharp change in BER as a function of SNR (a so-called cliff effect), which means that BER alone cannot be used as an early predictor of system failures. MER, on the other hand, can be used to measure signal-to-interference ratios accurately for such systems. Because MER

is a statistical measurement, its accuracy is directly related to the number of vectors, *N*, used in the computation. An accuracy of 0.14 dB can be obtained with *N* = 10,000, which would require about 2 ms to accumulate at the industry standard digital video rate of 5.057 Msymbols/s.

In order to design a proper optical data link, the contribution of different types of noise sources should be assessed when developing a link budget. There are two basic approaches to link budget modeling. One method is to design the link to operate at the desired BER when all the individual link components assume their worst case performance. This conservative approach is desirable when very high performance is required, or when it is difficult or inconvenient to replace failing components near the end of their useful lifetimes. The resulting design has a high safety margin; in some cases, it may be overdesigned for the required level of performance. Since it is very unlikely that all the elements of the link will assume their worst case performance at the same time, an alternative is to model the link budget statistically. For this method, distributions of transmitter power output, receiver sensitivity, and other parameters are either measured or estimated. They are then combined statistically using an approach such as the Monte Carlo method, in which many possible link combinations are simulated to generate an overall distribution of the available link optical power. A typical approach is the 3-sigma design, in which the combined variations of all link components are not allowed to extend more than 3 standard deviations from the average performance target in either direction. The statistical approach results in greater design flexibility and in generally increased distance compared with a worst case model at the same BER.

12.2. LINK BUDGET ANALYSIS: INSTALLATION LOSS

It is convenient to break down the link budget into two areas: installation loss and available power. Installation or DC loss refers to optical losses associated with the fiber cable plant, such as connector loss, splice loss, and bandwidth considerations. Available optical power is the difference between the transmitter output and receiver input powers, minus additional losses due to optical noise sources on the link (also known as AC losses). With this approach, the installation loss budget may be treated statistically and the available power budget, as worst case. First, we consider the installation loss budget, which can be broken down into three areas: transmission loss, fiber attenuation as a function of wavelength, and connector or splice losses.

12.2.1. Transmission Loss

Transmission loss is perhaps the most important property of an optical fiber; it affects the link budget and maximum unrepeated distance. Since the maximum optical power launched into an optical fiber is determined by international laser

eye safety standards [8], the number and separation between optical repeaters and regenerators are largely determined by this loss. The mechanisms responsible for this loss include material absorption as well as both linear and nonlinear scattering of light from impurities in the fiber [1–5]. Typical loss for single-mode optical fibers is about 2 to 3 dB/km near 800-nm wavelength, 0.5 dB/km near 1300 nm, and 0.25 dB/km near 1550 nm. Multimode fiber loss is slightly higher, and bending loss will only increase the link attenuation further.

12.2.2. Attenuation vs. Wavelength

Since fiber loss varies with wavelength, changes in the source wavelength or use of sources with a spectrum of wavelengths will produce additional loss. Transmission loss is minimized near the 1550-nm wavelength band, which unfortunately does not correspond with the dispersion minimum at around 1310 nm. An accurate model for fiber loss as a function of wavelength has been developed by Walker [9]; this model accounts for the effects of linear scattering, macrobending, and material absorption due to ultraviolet and infrared band edges, hydroxide [OH] absorption, and absorption from common impurities such as phosphorous. Using this model, it is possible to calculate the fiber loss as a function of wavelength for different impurity levels; the fiber properties can be specified along with the acceptable wavelength limits of the source to limit the fiber loss over the entire operating wavelength range. Design tradeoffs are possible between center wavelength and fiber composition to achieve the desired result. Typical loss due to wavelength dependent attenuation for laser sources on single-mode fiber can be held below 0.1 dB/km.

12.2.3. Connector and Splice Losses

There are also installation losses associated with fiber-optic connectors and splices; both of these are inherently statistical in nature and can be characterized by a Gaussian distribution. There are many different kinds of standardized optical connectors, some of which have been discussed previously. Some industry standards also specify the type of optical fiber and connectors suitable for a given application [10]. There are also different models which have been published for estimating connection loss due to fiber misalignment [11, 12]. Most of these models treat loss due to misalignment of fiber cores, offset of fibers on either side of the connector, and angular misalignment of fibers. The loss due to these effects is then combined into an overall estimate of the connector performance. No general model is available to treat all types of connectors, but typical connector loss values average about 0.5 dB worst case for multimode, slightly higher for single mode (see Table 12.1).

Optical splices are required for longer links, since fiber is usually available in spools of 1 to 5 km, or to repair broken fibers. There are two basic types:

Table 12.1
Datacom vs Telecom Requirements.

	Datacom	Telecom
BER	10e–12 to 10e–15	10e–9e
Distance	20–50 km	Varies with repeaters
No. transceivers/km	Large	Small
Signal bandwidth	00 Mb–1 Gb	3–5 Kb
Field service	Untrained users	Trained staff
No. fiber replugs	250–500	<100 over lifetime

mechanical splices (which involve placing the two fiber ends in a receptacle that holds them close together, usually with epoxy) and the more commonly used fusion splices (in which the fibers are aligned and then heated sufficiently to fuse the two ends together).

If the optical fiber is improperly cabled or installed, additional loss can be experienced due to bending of the fibers. This falls into two categories: microbending (due to nanometer-scale variations in the fiber) and macrobending (due to much larger, visible bends in the fiber). Because both types of bending loss may contribute to the attenuation of a single-mode or multimode fiber, it is obviously desirable to minimize bending in the application whenever possible. Most qualified optical fiber cables specify a maximum bend radius to limit macrobending effects, typically around 10–15 mm, although this varies with the cable type and manufacturer's recommendations.

12.3. LINK BUDGET ANALYSIS: OPTICAL POWER PENALTIES

Next, we will consider the assembly loss budget, which is the difference between the transmitter output and receiver input powers, allowing for optical power penalties due to noise sources in the link. We will follow the standard convention in the literature of assuming a digital optical communication link that is best characterized by its BER. Contributing factors to link performance include the following:

- Dispersion (modal and chromatic) or intersymbol interference
- Mode partition noise
- Mode hopping
- Extinction ratio
- Multipath interference
- Relative intensity noise (RIN)
- Timing jitter

- Radiation-induced darkening
- Modal noise

Higher order, nonlinear effects, including Stimulated Raman and Brillouin scattering and frequency chirping, will also be discussed.

12.3.1. Dispersion

The most important fiber characteristic after transmission loss is dispersion, or intersymbol interference. This refers to the broadening of optical pulses as they propagate along the fiber. As pulses broaden, they tend to interfere with adjacent pulses; this limits the maximum achievable data rate. In multimode fibers, there are two dominant kinds of dispersion: modal and chromatic. Modal dispersion refers to the fact that different modes will travel at different velocities and cause pulse broadening. The fiber's modal bandwidth in units of MHz-km is specified according to the expression

$$\mathrm{BW}_{\mathrm{modal}} = \mathrm{BW}_1/\mathrm{L}^{\gamma} \tag{12.4}$$

where BW_{modal} is the modal bandwidth for a length L of fiber, BW_1 is the manufacturer-specified modal bandwidth of a 1-km section of fiber, and γ is a constant known as the modal bandwidth concatenation length scaling factor. The term γ usually assumes a value between 0.5 and 1, depending on details of the fiber manufacturing and design as well as the operating wavelength; it is conservative to take $\gamma = 1.0$. Modal bandwidth can be increased by mode mixing, which promotes the interchange of energy between modes to average out the effects of modal dispersion. Fiber splices tend to increase the modal bandwidth, although it is conservative to discard this effect when designing a link.

The other major contribution is chromatic dispersion, BW_{chrom}, which occurs because different wavelengths of light propagate at different velocities in the fiber. For multimode fiber, this is given by an empirical model of the form

$$\mathrm{BW}_{\mathrm{chrom}} = \frac{\mathrm{L}^{\gamma c}}{\sqrt{\lambda_{\mathrm{w}}}\,(\mathrm{a}_{\mathrm{o}} + \mathrm{a}_1|\lambda_{\mathrm{c}} - \lambda_{\mathrm{eff}}|)} \tag{12.5}$$

where L is the fiber length in km; λ_c is the center wavelength of the source in nm; λ_w is the source FWHM spectral width in nm; λ_c is the chromatic bandwidth length scaling coefficient, a constant; λ_{eff} is the effective wavelength, which combines the effects of the fiber zero-dispersion wavelength and spectral loss signature; and the constants a_1 and a_o are determined by a regression fit of measured data. From Ref. [13], the chromatic bandwidth for 62.5/125 micron fiber is empirically given by

$$\mathrm{BW}_{\mathrm{chrom}} = \frac{10^4\,\mathrm{L}^{-0.69}}{\sqrt{\lambda_{\mathrm{w}}}\,(1.1 + 0.0189|\lambda_{\mathrm{c}} - 1370|)} \tag{12.6}$$

For this expression, the center wavelength was 1335 nm, and λ_{eff} was chosen midway between λ_c and the water absorption peak at 1390 nm. Although λ_{eff} was estimated in this case, the expression still provides a good fit to the data. For 50/125 micron fiber, the expression becomes

$$BW_{chrom} = \frac{10^4 L^{-0.65}}{\sqrt{\lambda_w}(1.01 + 0.0177|\lambda_c - 1330|)} \tag{12.7}$$

For this case, λ_c was 1313 nm, and the chromatic bandwidth peaked at λ_{eff} = 1330 nm. Recall that this is only one possible model for fiber bandwidth [1]. The total bandwidth capacity of multimode fiber BW_t is obtained by combining the modal and chromatic dispersion contributions, according to

$$\frac{1}{BW_t^2} = \frac{1}{BW_{chrom}^2} + \frac{1}{BW_{modal}^2} \tag{12.8}$$

Once the total bandwidth is known, the dispersion penalty can be calculated for a given data rate. One expression for the dispersion penalty in dB is

$$P_d = 1.22\left[\frac{\text{Bit Rate (Mb/s)}}{BW_t\text{(MHz)}}\right]^2 \tag{12.9}$$

For typical telecommunication grade fiber, the dispersion penalty for a 20-km link is about 0.5 dB.

Dispersion is usually minimized at wavelengths near 1310 nm; special types of fiber have been developed that manipulate the index profile across the core to achieve minimal dispersion near 1550 nm, which is also the wavelength region of minimal transmission loss. Unfortunately, this dispersion-shifted fiber suffers from some practial drawbacks, including susceptibility to certain kinds of nonlinear noise and increased interference between adjacent channels in a wavelength multiplexing environment. There is a new type of fiber that minimizes dispersion while reducing the unwanted crosstalk effects, called dispersion optimized fiber. By using a very sophisticated fiber profile, it is possible to minimize dispersion over the entire wavelength range from 1300 to 1550 nm, at the expense of very high loss (around 2 dB/km); this is known as dispersion flattened fiber. Yet another approach is called dispersion conpensating fiber; this fiber is designed with negative dispersion characteristics, so that when used in series with conventional fiber it will offset the normal fiber dispersion. Dispersion compensating fiber has a much narrower core than standard single-mode fiber, which makes it susceptible to nonlinear effects; it is also birefringent and suffers from polarization mode dispersion, in which different states of polarized light propagate with very different group velocities. Note that standard single-mode fiber does not preserve the polarization state of the incident light. There is yet another type of specialty fiber,

with asymmetric core profiles, capable of preserving the polarization of incident light over long distances.

By definition, single-mode fiber does not suffer modal dispersion. Chromatic dispersion is an important effect, though, even given the relatively narrow spectral width of most laser diodes. The dispersion of single-mode fiber corresponds to the first derivative of group velocity τ_g with respect to wavelength and is given by

$$D = \frac{d\tau_g}{d\lambda} = \frac{S_o}{4}\left(\lambda_c - \frac{\lambda_o^4}{\lambda_c^3}\right) \tag{12.10}$$

where D is the dispersion in ps/(km-nm) and λ_c is the laser center wavelength. The fiber is characterized by its zero-dispersion wavelength, λ_o, and zero-dispersion slope, S_o. Usually, both center wavelength and zero-dispersion wavelength are specified over a range of values; it is necessary to consider both upper and lower bounds in order to determine the worst case dispersion penalty. This can be seen from Fig. 12.3, which plots D vs. wavelength for some typical values of λ_o and λ_c; the largest absolute value of D occurs at the extremes of this region. Once the dispersion is determined, the intersymbol interference penalty as a function of link length, L, can be determined to a good approximation from a model proposed by Agrawal [14]:

$$P_d = 5\log(1 + 2\pi(BD\,\Delta\lambda)^2\,L^2) \tag{12.11}$$

where B is the bit rate and $\Delta\lambda$ is the root mean square (RMS) spectral width of the source. By maintaining a close match between the operating and zero-dispersion wavelengths, this penalty can be kept to a tolerable 0.5–1.0 dB in most cases.

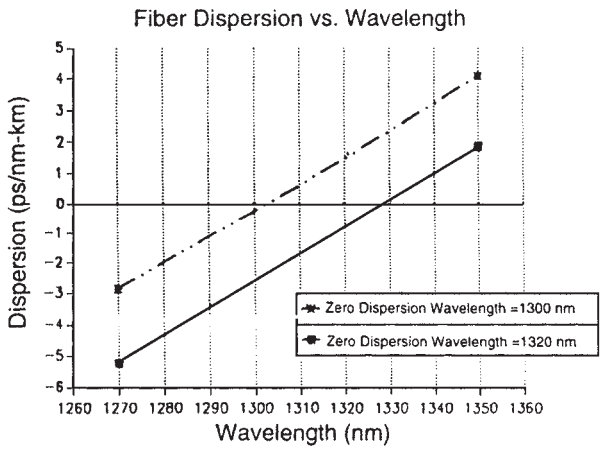


Figure 12.3 Single-mode fiber dispersion as a function of wavelength.

12.3.2. Mode Partition Noise

Group velocity dispersion contributes to another optical penalty that remains the subject of continuing research, mode partition noise and mode hopping. This penalty is related to the properties of a Fabry-Perot type laser diode cavity; although the total optical power output from the laser may remain constant, the optical power distribution among the laser's longitudinal modes will fluctuate. This is illustrated by the model depicted in Fig. 12.4. When a laser diode is directly modulated with injection current, the total output power stays constant

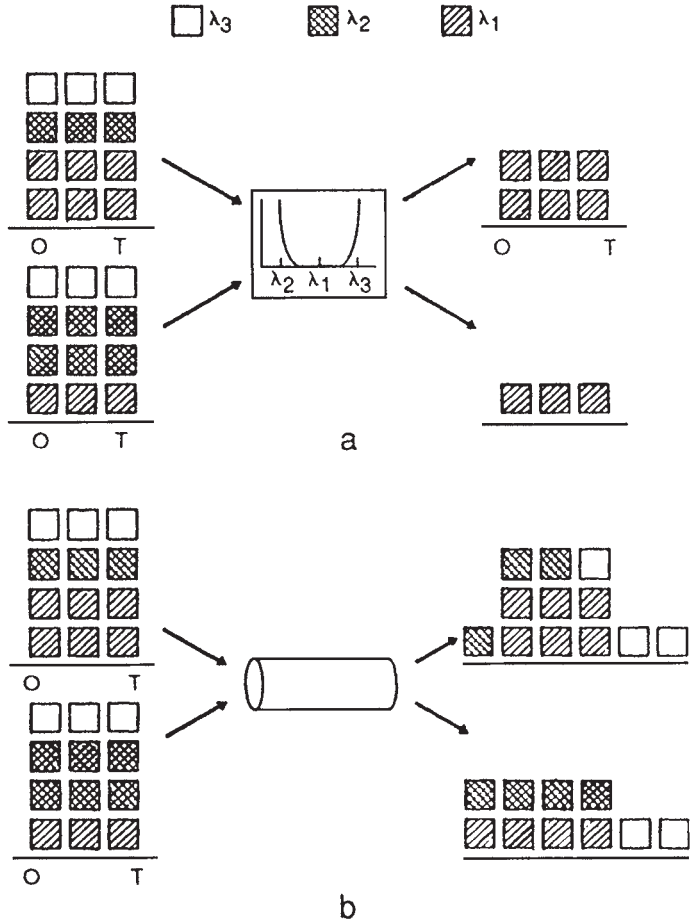


Figure 12.4 Model for mode partition noise; an optical source emits a combination of wavelengths, illustrated by different color blocks: (a) wavelength-dependent loss; (b) chromatic dispersion.

from pulse to pulse; however, the power distribution among several longitudinal modes will vary between pulses. We must be careful to distinguish this behavior of the instantaneous laser spectrum, which varies with time, from the time-averaged spectrum that is normally observed experimentally. The light propagates through a fiber with wavelength-dependent dispersion or attenuation, which deforms the pulse shape. Each mode is delayed by a different amount due to group velocity dispersion in the fiber; this leads to additional signal degradation at the receiver, in addition to the intersymbol interference caused by chromatic dispersion alone, discussed earlier. This is known as mode partition noise; it is capable of generating bit error rate floors, such that additional optical power into the receiver will not improve the link BER. This is because mode partition noise is a function of the laser spectral fluctuations and wavelength-dependent dispersion of the fiber, so the signal-to-noise ratio due to this effect is independent of the signal power. The power penalty due to mode partition noise was first calculated by Ogawa [15] as

$$P_{mp} = 5\log\frac{1}{(1-Q^2\sigma_{mp}^2)} \tag{12.12}$$

where

$$\sigma_{mp}^2 = \frac{1}{2}k^2(\pi B)^4[A_1^4\Delta\lambda^4 + 42A_1^2A_2^2\Delta\lambda^6 + 48A_2^4\Delta\lambda^8] \tag{12.13}$$

$$A_1 = DL \tag{12.14}$$

and

$$A_2 = \frac{A_1}{2(\lambda_c - \lambda_o)} \tag{12.15}$$

The mode partition coefficient, *k*, is a number between 0 and 1, which describes how much of the optical power is randomly shared between modes; it summarizes the statistical nature of mode partition noise. According to Ogawa, *k* depends on the number of interacting modes and rms spectral width of the source, the exact dependence being complex. However, subsequent work has shown [16] that Ogawa's model tends to underestimate the power penalty due to mode partition noise because it does not consider the variation of longitudinal mode power between successive baud periods, and because it assumes a linear model of chromatic dispersion rather than the nonlinear model given in the above equation. A more detailed model has been proposed by Campbell [17], which is general enough to include effects of the laser diode spectrum, pulse shaping, transmitter extinction ratio, and statistics of the datastream. While Ogawa's model assumed an equiprobable distribution of zeros and ones in the data stream, Campbell showed that mode partition noise is data dependent as well. Recent work based on this model [18] has re-derived the signal variance:

$$\sigma_{mp}^2 = E_{av}(\sigma_o^2 + \sigma_{+1}^2 + \sigma_{-1}^2) \tag{12.16}$$

where the mode partition noise contributed by adjacent baud periods is defined by

$$\sigma_{+1}^2 + \sigma_{-1}^2 = \frac{1}{2}k^2(\pi B)^4[1.25A_1^4\Delta\lambda^4 + 40.95A_1^2A_2^2\Delta\lambda^6 + 50.25A_2^4\Delta\lambda^8] \tag{12.17}$$

and the time-average extinction ratio $E_{av} = 10\log(P_1/P_0)$, where P_1, P_0 represent the optical power by a "1" and "0", respectively. If the operating wavelength is far away from the zero-dispersion wavelength, the noise variance simplifies to

$$\sigma_{mp}^2 = 2.25\frac{k^2}{2}E_{av}(1 - e^{-\beta L^2})^2 \tag{12.18}$$

which is valid provided that

$$\beta = (\pi BD\ \Delta\lambda)^2 << 1 \tag{12.19}$$

Many diode lasers exhibit mode hopping or mode splitting in which the spectrum appears to split optical power between two or three modes for brief periods of time. The exact mechanism is not fully understood, but stable Gaussian spectra are generally only observed for CW operation and temperature stabilized lasers. During these mode hops, the above theory does not apply since the spectrum is non-Gaussian, and the model will overpredict the power penalty. Hence, it is not possible to model mode hops as mode partitioning with $k = 1$. There is no currently published model describing a treatment of mode hopping noise, although recent papers [19] suggest approximate calculations based on the statistical properties of the laser cavity. In a practical link, some amount of mode hopping is probably unavoidable as a contributor to burst noise; empirical testing of link hardware remains the only reliable way to reduce this effect. A practical rule of thumb is to keep the mode partition noise penalty less than 1.0 dB maximum, provided that this penalty is far away from any noise floors.

12.3.3. Extinction Ratio

The receiver extinction ratio also contributes directly to the link penalties. The receiver BER is a function of the modulated AC signal power; if the laser transmitter has a small extinction ratio, the DC component of total optical power is significant. Gain or loss can be introduced in the link budget if the extinction ratio at which the receiver sensitivity is measured differs from the worst case transmitter extinction ratio. If the extinction ratio E_t at the transmitter is defined as the ratio of optical power when a one is transmitted vs. when a zero is transmitted,

$$E_t = \frac{\text{Power(1)}}{\text{Power(0)}} \tag{12.20}$$

then we can define a modulation index at the transmitter M_t according to

$$M_t = \frac{E_t - 1}{E_t + 1} \tag{12.21}$$

Similarly, we can measure the linear extinction ratio at the optical receiver input and define a modulation index M_r. The extinction ratio penalty is given by

$$P_{er} = -10 \log\left(\frac{M_t}{M_r}\right) \tag{12.22}$$

where the subscripts T and R refer to specifications for the transmitter and receiver, respectively. Usually, the extinction ratio is specified to be the same at the transmitter and receiver and is large enough so that there is no power penalty due to extinction ratio effects.

Furthermore, the extinction ratio is used to calculate the optical modulation amplitude (OMA), which is sometimes specified in place of receiver sensitivity (for example, in the ANSI Fibre Channel Standard recent revisions). OMA is defined as the difference in optical power between logic levels of 1 and 0; in terms of average optical power (in microwatts) and extinction ratio, it is given by

$$OMA = 2P_{av}\left(\frac{E - 1}{E + 1}\right) \tag{12.23}$$

where the extinction ratio in this case is the absolute (unitless linear) ratio of average optical power (in microwatts) between a logic level 1 and 0, measured under fully modulated conditions in the presence of worst case reflections. In the Fibre Channel Standard, for example, the OMA specified at 1.0625 Gbit/s for short-wavelength (850-nm) laser sources is between 31 and 2000 microwatts (peak-to-peak), which is equivalent to an average power of −17 dBm and an extinction ratio of 9 dB. Similarly, the OMA specified at 2.125 Gbit/s for short-wavelength (850 nm) laser sources is between 49 and 2000 microwatts (peak-to-peak), which is equivalent to an average power of −15 dBm and an extinction ratio of 9 dB.

12.3.4. Multipath Interference

Another important property of the optical link is the amount of reflected light from the fiber end faces that return up the link back into the transmitter. Whenever there is a connection or splice in the link, some fraction of the light is reflected back; each connection is thus a potential noise generator, since the reflected fields can interfere with one another to create noise in the detected optical signal. The phenomenon is analogous to the noise caused by multiple atmospheric reflections of radio waves and is known as multipath interference noise. To limit this noise, connectors and splices are specified with a minimum return loss. If there are a total of N reflection points in a link and the geometric mean of the connector

reflections is alpha, then based on the model of Duff et al. [20] the power penalty due to multipath interference (adjusted for bit error rate and bandwidth) is closely approximated by

$$P_{mpi} = 10\log(1 - 0.7Na) \tag{12.24}$$

Multipath noise can usually be reduced well below 0.5 dB with available connectors, whose return loss is often better than 25 dB.

12.3.5. Relative Intensity Noise (RIN)

Stray light reflected back into a Fabry-Perot type laser diode gives rise to intensity fluctuations in the laser output. This is a complicated phenomenon, strongly dependent on the type of laser; it is called either reflection-induced intensity noise or relative intensity noise (RIN). This effect is important because it can also generate BER floors. The power penalty due to RIN is the subject of ongoing research; since the reflected light is measured at a specified signal level, RIN is data dependent, although it is independent of link length. Since many laser diodes are packaged in windowed containers, it is difficult to correlate the RIN measurements on an unpackaged laser with those of a commercial product. Several detailed attempts have been made to characterize RIN [21, 22]. Typically, the RIN noise is assumed Gaussian in amplitude and uniform in frequency over the receiver bandwidth of interest. The RIN value is specified for a given laser by measuring changes in the optical power when a controlled amount of light is fed back into the laser. It is signal dependent, and it is also influenced by temperature, bias voltage, laser structure, and other factors that typically influence laser output power [22]. If we assume that the effect of RIN is to produce an equivalent noise current at the receiver, then the additional receiver noise σ_r may be modeled as

$$\sigma_r = \gamma^2 S^{2g} B \tag{12.25}$$

where S is the signal level during a bit period, B is the bit rate, and g is a noise exponent that defines the amount of signal-dependent noise. If $g = 0$, noise power is independent of the signal, while for $g = 1$, noise power is proportional to the square of the signal strength. The coefficient γ is given by

$$\gamma^2 = S_i^{2(1-g)} 10^{(RIN_i/10)} \tag{12.26}$$

where RIN_i is the measured RIN value at the average signal level S_i, including worst case backreflection conditions and operating temperatures. The Gaussian BER probability due to the additional RIN noise current is given by

$$P_{error} = \frac{1}{2}\left[P_e^1\left(\frac{S_1 - S_o}{2\sigma_1} \right) + P_e^0\left(\frac{S_1 - S_o}{2\sigma_o} \right) \right] \tag{12.27}$$

where σ_1, σ_o represent the total noise current during transmission of a digital 1 and 0, respectively, and P_e^1, P_e^o are the probabilities of error during transmission of a 1 or 0, respectively. The power penalty due to RIN may then be calculated by determining the additional signal power required to achieve the same BER with RIN noise present as without the RIN contribution. One approximation for the RIN power penalty is given by

$$P_{rin} = -5\log\left[1 - Q^2(BW)(1+M_r)^{2g}(10^{RIN/10})\left(\frac{1}{M_r}\right)^2\right] \qquad (12.28)$$

where the RIN value is specified in dB/Hz, BW is the receiver bandwidth, M_r is the receiver modulation index, and the exponent g is a constant varying between 0 and 1, which relates the magnitude of RIN noise to the optical power level. The maximum RIN noise penalty in a link can usually be kept to below 0.5 dB.

12.3.6. Jitter

Although it is not strictly an optical phenomenon, another important area in link design deals with the effects of timing jitter on the optical signal. In a typical optical link, a clock is extracted from the incoming data signal, which is used to retime and reshape the received digital pulse. The received pulse is then compared with a threshold to determine if a digital 1 or 0 was transmitted. So far, we have discussed BER testing with the implicit assumption that the measurement was made in the center of the received data bit; to achieve this, a clock transition at the center of the bit is required. When the clock is generated from a receiver timing recovery circuit, it will have some variation in time and the exact location of the clock edge will be uncertain. Even if the clock is positioned at the center of the bit, its position may drift over time.

There will be a region of the bit interval, or eye, in the time domain where the BER is acceptable; this region is defined as the eyewidth [1–3]. Eyewidth measurements are an important parameter for evaluation of fiber-optic links; they are intimately related to the BER, as well as the acceptable clock drift, pulsewidth distortion, and optical power. At low optical power levels, the receiver signal-to-noise ratio is reduced; increased noise causes amplitude variations in the received signal. These amplitude variations are translated into time domain variations in the receiver decision circuitry, which narrows the eyewidth. At the other extreme, an optical receiver may become saturated at high optical power, reducing the eyewidth and making the system more sensitive to timing jitter. This behavior results in the typical "bathtub" curve shown in Fig. 12.2. For this measurement, the clock is delayed from one end of the bit cell to the other, with the BER calculated at each position. Near the ends of the cell, a large number of errors occur; toward the center of the cell, the BER decreases to its true value. The eye opening may be defined as the portion of the eye for which the BER remains constant;

pulsewidth distortion occurs near the edges of the eye, which denotes the limits of the valid clock timing. Uncertainty in the data pulse arrival times causes errors to occur by closing the eye window and causing the eye pattern to be sampled away from the center. This is one of the fundamental problems of optical and digital signal processing, and a large body of work has been done in this area [23, 24]. In general, multiple jitter sources will be present in a link; these will tend to be uncorrelated. However, jitter on digital signals, especially resulting from a cascade of repeaters, may be coherent.

International standards on jitter were first published by the CCITT (Central Commission for International Telephony and Telegraphy, now known as the International Telecommunications Union, or ITU). This standards body has adopted a definition of jitter [24] as short-term variations of the significant instants (rising or falling edges) of a digital signal from their ideal position in time. Longer-term variations are described as wander; in terms of frequency, the distinction between jitter and wander is somewhat unclear. The predominant sources of jitter include the following:

- Phase noise in receiver clock recovery circuits, particularly crystal-controlled oscillator circuits; this may be aggravated by filters or other components that do not have a linear phase response. Noise in digital logic resulting from restricted rise and fall times may also contribute to jitter.
- Imperfect timing recovery in digital regenerative repeaters, which is usually dependent on the data pattern.
- Different data patterns, which may contribute to jitter when the clock recovery circuit of a repeater attempts to recover the receive clock from inbound data. Data pattern sensitivity can produce as much as 0.5 dB penalty in receiver sensitivity. Higher data rates are more susceptible (>1 Gbit/s); data patterns with long run lengths of 1's or 0's, or with abrupt phase transitions between consecutive blocks of 1's and 0's, tend to produce worst case jitter.
- At low optical power levels, the receiver signal-to-noise ratio, Q, is reduced; increased noise causes amplitude variations in the signal, which may be translated into time domain variations by the receiver circuitry.
- Low frequency jitter, also called wander, resulting from instabilities in clock sources and modulation of transmitters.
- Very-low-frequency jitter caused by variations in the propagation delay of fibers, connectors, and the like, typically resulting from small temperature variations (making it especially difficult to perform long-term jitter measurements).

In general, jitter from each of these sources will be uncorrelated; jitter related to modulation components of the digital signal may be coherent, and cumulative

jitter from a series of repeaters or regenerators may also contain some well-correlated components.

There are several parameters of interest in characterizing jitter performance. Jitter may be classified as either random or deterministic, depending on whether it is associated with pattern dependent effects. These are distinct from the duty cycle distortion that often accompanies imperfect signal timing. Each component of the optical link (data source, serializer, transmitter, encoder, fiber, receiver, retiming/clock recovery/deserialization, decision circuit) will contribute some fraction of the total system jitter. If we consider the link to be a "black box" (but not necessarily a linear system), then we can measure the level of output jitter in the absence of input jitter; this is known as the "intrinsic jitter" of the link. The relative importance of jitter from different sources may be evaluated by measuring the spectral density of the jitter. Another approach is the maximum tolerable input jitter (MTIJ) for the link. Finally, since jitter is essentially a stochastic process, we may attempt to characterize the jitter transfer function (JTF) of the link, or estimate the probability density function of the jitter. When multiple traces occur at the edges of the eye, this can indicate the presence of data-dependent jitter or duty cycle distortion; a histogram of the edge location will show several distinct peaks. This type of jitter can indicate a design flaw in the transmitter or receiver. By contrast, random jitter typically has a more Gaussian profile and is present to some degree in all data links.

The problem of jitter accumulation in a chain of repeaters becomes increasingly complex; however, we can state some general rules of thumb. It has been shown [25] that jitter can be generally divided into two components, one due to repetitive patterns and one due to random data. In receivers with phase-locked loop timing recovery circuits, repetitive data patterns will tend to cause jitter accumulation, especially for long run lengths. This effect is commonly modeled as a second-order receiver transfer function. Jitter will also accumulate when the link is transferring random data. Jitter due to random data is of two types: systematic and random. The classic model for systematic jitter accumulation in cascaded repeaters was published by Byrne [26]. The Byrne model assumes cascaded identical timing recovery circuits, and then the systematic and random jitter can be combined as rms quantities so that total jitter due to random jitter may be obtained. This model has been generalized to networks consisting of different components [27] and to nonidentical repeaters [28]. Despite these considerations, for well-designed practical networks the basic results of the Byrne model remain valid for N nominally identical repeaters transmitting random data; systematic jitter accumulates in proportion to $N^{1/2}$ and random jitter accumulates in proportion to $N^{1/4}$. For most applications, the maximum timing jitter should be kept below about 30% of the maximum receiver eye opening.

12.3.7. Modal Noise

An additional effect of lossy connectors and splices is modal noise. Because high-capacity optical links tend to use highly coherent laser transmitters, random coupling between fiber modes causes fluctuations in the optical power coupled through splices and connectors; this phenomenon is known as modal noise [29]. As one might expect, modal noise is worst when using laser sources in conjunction with multimode fiber; recent industry standards have allowed the use of short-wave lasers (750–850 nm) on 50-micron fiber that may experience this problem. Modal noise is usually considered to be nonexistent in single-mode systems. However, modal noise in single-mode fibers can arise when higher order modes are generated at imperfect connections or splices. If the lossy mode is not completely attenuated before it reaches the next connection, interference with the dominant mode may occur. The effects of modal noise have been modeled previously [29], assuming that the only significant interaction occurs between the LP01 and LP11 modes for a sufficiently coherent laser. For N sections of fiber, each of length L in a single-mode link, the worst case sigma for modal noise can be given by

$$\sigma_m = \sqrt{2}N\eta(1-\eta)e^{-aL} \tag{12.29}$$

where α is the attenuation coefficient of the LP_{11} mode, and η is the splice transmission efficiency, given by

$$\eta = 10^{-(\eta_o/10)} \tag{12.30}$$

where η_o is the mean splice loss (typically, splice transmission efficiency will exceed 90%). The corresponding optical power penalty due to modal noise is given by

$$P = -5\log(1-Q^2\sigma_m^2) \tag{12.31}$$

where Q corresponds to the desired BER. This power penalty should be kept to less than 0.5 dB.

12.3.8. Radiation-Induced Loss

Another important environmental factor, as mentioned earlier, is exposure of the fiber to ionizing radiation damage. There is a large body of literature concerning the effects of ionizing radiation on fiber links [30, 31]. Many factors can affect the radiation susceptibility of optical fiber, including the type of fiber, type of radiation (gamma radiation is usually assumed to be representative), total dose, dose rate (important only for higher exposure levels), prior irradiation history of the fiber, temperature, wavelength, and data rate. Optical fiber with a pure silica core is least susceptible to radiation damage. However, almost all commercial fiber is intentionally doped to control the refractive index of the core and cladding, as

well as dispersion properties. Trace impurities are also introduced, which become important only under irradiation. Among the most important are Ge dopants in the core of graded-index (GRIN) fibers, in addition to F, Cl, P, B, OH content, and the alkali metals. In general, radiation sensitivity is worst at lower temperatures and is also made worse by hydrogen diffusion from materials in the fiber cladding. Because of the many factors involved, there does not exist a comprehensive theory to model radiation damage in optical fibers.

The basic physics of the interaction has been described [30, 31]; there are two dominant mechanisms, radiation-induced darkening and scintillation. First, high-energy radiation can interact with dopants, impurities, or defects in the glass structure to produce color centers that absorb strongly at the operating wavelength. Carriers can also be freed by radiolytic or photochemical processes. Some of these become trapped at defect sites, which modifies the band structure of the fiber and causes strong absorption at infrared wavelengths. This radiation-induced darkening increases the fiber attenuation. In some cases, it is partially reversible when the radiation is removed, although high levels or prolonged exposure will permanently damage the fiber. A second effect is caused if the radiation interacts with impurities to produce stray light, or scintillation. This light is generally broadband, but will tend to degrade the BER at the receiver; scintillation is a weaker effect than radiation-induced darkening. These effects will degrade the BER of a link; they can be prevented by shielding the fiber, or can be partially overcome by a third mechanism, photobleaching. The presence of intense light at the proper wavelength can partially reverse the effects of darkening in a fiber. It is also possible to treat silica core fibers by briefly exposing them to controlled levels of radiation at controlled temperatures; this increases the fiber loss, but makes the fiber less susceptible to future irradiation. These so-called radiation hardened fibers are often used in environments where radiation is anticipated to play an important role. Recently, several models have been advanced [31] for the performance of fiber under moderate radiation levels; the effect on BER is a power law model of the form

$$\mathrm{BER} = \mathrm{BER}_o + A(\mathrm{dose})^b \tag{12.32}$$

where BER_0 is the link BER prior to irradiation, the dose is given in rads, and the constants A and *b* are empirically fitted. The loss due to normal background radiation exposure over a typical link lifetime can be held below about 0.5 dB.

12.3.9. Nonlinear Noise Effects

Except for those penalties that produce BER floors such as mode partitioning and RIN, most penalties can be reduced by increasing the transmitted or received optical power. This brute-force approach is subject to limitations such as maintaining class 1 laser safety; even if it were possible to increase optical power

significantly (as will be discussed later using optical amplifiers), a new class of nonlinear optical penalties becomes important at high-power levels. (Some of these are discussed here; the reader is referred to Ref. [54] for a more complete treatment of nonlinear phenomenon.) Class 1 laser systems typically do not experience these nonlinear effects, although they may be important for optical fiber amplifiers or systems using open fiber control (OFC), for which significantly higher power levels may be present in the fiber.

At higher-optical power levels, nonlinear scattering may limit the behavior of a fiber-optic link. The dominant effects are stimulated Raman and Brillouin scattering. When incident optical power exceeds a threshold value, significant amounts of light may be scattered from small imperfections in the fiber core or by mechanical (acoustic) vibrations in the transmission media. These vibrations can be caused by the high-intensity electromagnetic fields of light concentrated in the core of a single-mode fiber. Because the scattering process also involves the generation of photons, the scattered light can be frequency shifted [55]. Put another way, we can think of the high-intensity light as generating a regular pattern of very slight differences in the fiber refractive index. This creates a moving diffraction grating in the fiber core, and the scattered light from this grating is Doppler shifted in frequency by about 11 GHz. This effect is known as stimulated Brillouin scattering (SBS); under these conditions, the output light intensity becomes nonlinear as well. Stimulated Brillouin scattering will not occur below the optical power threshold defined by

$$Pc = 21A/G_bL_c \text{ watts} \tag{12.33}$$

where L_c is the effective interaction length, A is the cross-sectional area of the guided mode, and G_b is the Brillouin gain coefficient [55]. Brillouin scattering has been observed in single-mode fibers at wavelengths greater than cutoff with optical power as low as 5 mW. It can be a serious problem in long-distance communication systems when the span between amplifiers is low and the bit rate is less than about 2 Gbit/s, in WDM systems up to about 10 Gbit/s when the spectral width of the signal is very narrow, or in remote pumping of some types of optical amplifiers. In general, SBS is worse for narrow laser linewidths (and is generally not a problem for channel bandwidth greater than 100 MHz), wavelength (SBS is worse near 1550 nm than near 1300 nm), and signal power per unit area in the fiber core. All of these factors are summarized in the above expression.

SBS can be a concern in long-distance communication systems, when the span between amplifiers is large and the bit rate is below about 2.5 Gbit/s, in WDM systems where the spectral width of the signal is very narrow, and in remote pumping of optical amplifiers using narrow linewidth sources. In cases where SBS could be a problem, the source linewidth can be intentionally broadened by using an external modulator or additional RF modulation on the laser injection current. However, this is a tradeoff against long-distance transmission,

as broadening the linewidth also increases the effects of chromatic dispersion. When the scattered light experiences frequency shifts outside the acoustic phonon range, due instead to modulation by impurities or molecular vibrations in the fiber core, the effect is known as stimulated Raman scattering (SRS). The mechanism is similar to SBS**,** and scattered light can occur in both forward and backward directions along the fiber. The threshold below which Raman scattering will not occur is given by

$$P_t = 16A/G_R L_c \text{ watts} \tag{12.34}$$

where G_R is the Raman gain coefficient [54]. Note that SRS is also influenced by fiber dispersion and that standard fiber reduces the effect of SRS by half (3 dB) compared with dispersion-shifted fiber.

As a rule of thumb, the optical power threshold for Raman scattering is about three times larger than for Brillouin scattering. Another good rule of thumb is that SRS can be kept to acceptable levels if the product of total power and total optical bandwidth is less than 500 GHz-W. This is quite a lot; for example, consider a 10-channel DWDM system with standard wavelength spacing of 1.6 nm (200 GHz). The bandwidth becomes $200 \times 10 = 2000$ GHz, so the total power in all 10 channels would be limited to 250 mW in this case (in most DWDM systems, each channel will be well below 10 mW for other reasons such as laser safety considerations). In single-mode fiber, typical thresholds for Brillouin scattering are about 10 mW and for Raman scattering about 35 mW. These effects rarely occur in multimode fiber, where the thresholds are about 150 mW and 450 mW, respectivly. In general, the effect of SRS becomes greater as the signals are moved further apart in wavelength (within some limits); this introduces a tradeoff with FWM, which is reduced as the signal spacing increases. Optical amplifiers can be constructed using the principle of SRS. If a pump signal with relatively high power (half a watt or more) and a frequency 13.2 THz higher than the signal frequency is coupled into a sufficiently long length of fiber (about 1 km), then amplification of the signal will occur. Unfortunately, more efficient amplifiers require that the signal and pump wavelengths be spaced by almost exactly the Raman shift of 13.2 THz. Otherwise the amplification effect is greatly reduced.

It is not possible to build high-power lasers at arbitrary signal wavelengths; one possible solution is to build a pump laser at a convenient wavelength, and then wavelength shift the signal by the desired amount. However, another good alternative to SRS amplifiers are the widely used erbium doped fiber amplifiers (EDFAs). These allow the amplification of optical signals along their direction of travel in a fiber, without the need to convert back and forth from the electrical domain. While there are other types of optical amplifiers based on other rare-earth elements such as praseodymium (Pd) or neodymium (Nd), and even some optical amplifiers based on semiconductor devices, the erbium doped amplifiers are the most widely used because of their maturity and good performance at wavelengths

of interest near 1550 nm. Note that there is still a good deal of research going on in this area, most recently including thulium-doped fiber amplifiers (TDFAs) based on flurozirconiate (ZBLAN) fiber for use at wavelengths between 1450 and 1510 nm (the S-band).

The final nonlinear effect we will consider is frequency chirping of the optical signal. Chirping refers to a change in frequency with time. It takes its name from the sound of an acoustic signal whose frequency increases or decreases linearly with time. There are three ways in which chirping can affect a fiber-optic link. First, the laser transmitter can be chirped as a result of physical processes within the laser [56]; the effect has its origin in carrier-induced refractive index changes, making it an inevitable consequence of high-power direct modulation of semiconductor lasers. For lasers with low levels of relaxation oscillation damping, a model has been proposed for the chirped power penalty [56]:

$$P = 10\log\left(1+\frac{\pi B^2\lambda^2 LD\alpha}{4c}\right) \tag{12.35}$$

where c is the speed of light, B is the fiber bandwidth, λ is the wavelength of light, L is the length of the fiber, D is the dispersion, and α is the linewidth enhancement factor (a typical value is −4.5). This model is only a first approximation because it neglects the dependence of chirp on extinction ratio and nonlinear effects such as spectral hole burning [57]. Second, a sufficiently intense light pulse will be chirped by the nonlinear process of self-phase modulation in an optical fiber [58]**.** The effect arises from the interaction of the light and the intensity-dependent portion of the fiber's refractive index. It is thus dependent on the material and structure of the fiber, polarization of the light, and shape of the incident optical pulse. Based on a model from Ref. [58], the maximum optical power level before the spectral width increases by 2 nm is given by

$$P = \frac{n^2 A}{377 n_2 \kappa L_e} \text{ watts} \tag{12.36a}$$

where

$$L_e = (1/a_0)(1 - \exp(-a_0 L)) \tag{12.36b}$$

where n is the fiber's refractive index, κ is the propagation constant (a typical value is 7×10^4 at a wavelength of 1.3 μm) a_0 is the fiber attenuation coefficient, A is the fiber core cross-sectional area, and n_2 is the nonlinear coefficient of the fiber's refractive index (a typical value is 6.1×10^{-19} and L_e is the effective interaction length for the nonlinear interaction, which is related to the actual length of the fiber, L, by Eq. (12.36). This expression should be multipled by 2 if the fiber is not polarization preserving. Typically, this effect is not significant for optical power levels less than 950 mW in a single-mode fiber at 1.3-μm wavelength.

Finally, there is a power penalty arising from the propagation of a chirped optical pulse in a dispersive fiber because the new frequency components propagate at different group velocities. This may be treated as simply a much worse case of the conventional dispersion penalty [59], provided that one of the first two effects exists to chirp the optical signal.

12.4. GIGABIT ETHERNET LINK BUDGET MODEL

When the IEEE 802.3z committee began investigating high-data-rate Ethernet links in the late 1990s, a model was developed to predict the performance of laser-based multimode optical fiber links and potential tradeoffs between the various link penalties as part of preparing the link specifications [32, 33]. This model was first released to the public in 2001, has been enhanced periodically over the years, and now forms the basis for the gigabit Ethernet and 10 gigabit Ethernet specifications, as well as other multimode fiber standards including InfiniBand and several types of parallel optical links. (Note that this is not a transceiver design tool; rather it provides a framework for comparing different link design options.)

As this model has evolved over the years, there have been many updates, including the addition of single-mode specifications and polarization mode dispersion penalties, deterministic jitter, baseline wander, and other features. For the purposes of this chapter, we will review only the basic model, following the description in Ref. [32] (a spreadsheet implementation of the latest model with a change history is available from the IEEE [33]). The basic model is an extension of previously reported work on LED-based links [34, 35]. In addition to considering loss due to fiber attenuation, connectors, and splices, power penalties are calculated to account for the effects of intersymbol interference (ISI) [36], mode partition noise [37], extinction ratio, modal noise [38], and relative intensity noise (RIN). Since all power penalties except ISI are assumed to be independent of the link length, the ISI penalty dominates at longer distances and, along with link attenuation, determines the maximum link length for Ethernet applications.

This model assumes that the laser and multimode fiber impulse responses are Gaussian and that the optical receiver is nonequalized, with a single-pole filter having a specified 3-dB electrical bandwidth. Variations on this model have been published with results for the case of a nonequalized receiver having a raised cosine response [35] and for the case where the receiver has an exponential impulse response [34, 35].

The model includes expressions that convert the root mean square (rms) impulse response of the laser, fiber, and optical receiver to rise times, fall times, and bandwidths, which are used to determine the fiber and composite channel output impulse response and the ISI penalty. It is assumed that rise and fall times

are equal; to be conservative when modeling real components, the larger of the experimentally measured rise or fall time should be used.

First, we will consider the effects of rise/fall times, pulsewidth, and bandwidth; by calculating these factors, we can estimate the resulting eye diagrams for various link designs. Note that the Gigabit and 10 Gigabit Ethernet standards define two types of receiver sensitivity, namely, "normal" sensitivity (measured with a typical or good performance transmitter) and "stressed" sensitivity (measured with a transmitter having worst case jitter performance). While the stressed sensitivity is intended to prove interoperability, it is not required in order to use this link model. It has been shown [39] that for two positive signal pulses h1(t) and h2(t), if their convolution is represented by a signal h3(t) such that

$$h3(t) = h1(t)*h2(t) \tag{12.37}$$

then

$$\vartheta_3 = \sqrt{\vartheta_1^2 + \vartheta_2^2} \tag{12.38}$$

where σi is the rms pulsewidth of the individual components. The 10% to 90% rise time, T_i, and bandwidth of individual components, BW_i, are related by constant conversion factors, a_i and b_i, such that:

$$\vartheta_i(BW) = \frac{a_i}{BW_i} \tag{12.39}$$

and

$$\vartheta_i(T) = \frac{T_i}{b_i} \tag{12.40}$$

therefore

$$T_i = \frac{a_i b_i}{BW_i(6dB)} \tag{12.41}$$

where BW(6 dB) is the 6-dB electrical bandwidth (equivalent to the 3-dB optical bandwidth). Since equation 12.40 can be generalized to include the sum of an arbitrary number of signals, the rms pulsewidths of the individual components may be used to calculate the bandwidth or the 10% to 90% rise time of the composite system if the appropriate conversion factors for each individual component are known [35]. For example, the overall system rise time, T_s, may be calculated using:

$$T_s^2 = \sum_i \left(\frac{b_s T_i}{b_i} \right)^2 = \sum_i \left(\frac{a_i b_i}{BW_i(6dB)} \right)^2 = \sum_i \left(\frac{C_i}{BW_i(6dB)} \right)^2 \tag{12.42}$$

Using this approach and the central limit theorem, it can be shown that the composite impulse response of multimode fiber-optic links tend to a Gaussian impulse response [2]. For these systems, it can be shown that the conversion factors *a* and *b* are equal to 0.187 and 2.563, respectively, so that $C = 0.48$ [35]. These conversion factors can be applied to the laser and fiber.

The simplest form of optical receiver is a nonequalized receiver with a single-pole filter. This type of receiver can be modeled by an exponential impulse response of the form [35]:

$$\mathrm{hr(t)} = \frac{1}{\tau}\exp(-\mathrm{t}/\tau) \tag{12.43}$$

for t greater than or equal to zero; $hr(t)$ is zero otherwise. Here, τ is called the rise time constant; this impulse response has an rms width of τ. If the receiver is excited by a step function, then the 10% to 90% rise time of the source is [35]:

$$\mathrm{tr} = \ln(9).\tau \tag{12.44}$$

and the 3-dB bandwidth is [35]:

$$\mathrm{BW_r(3dB)} = 0.1588/\tau = 0.1588/\vartheta \tag{12.45}$$

By substitution we have

$$\mathrm{tr} = \ln(9)\left(\frac{0.1588}{\mathrm{BW_r(3dB)}}\right) \tag{12.46}$$

Therefore, $a = 0.1588$ and $b = \ln(9)$, for a component or system with an exponential impulse response. With the assumption that the fiber exit impulse response is Gaussian, we can calculate the fiber 10% to 90% exit response time (Te):

$$\mathrm{T_{exit}} = \sqrt{\left(\frac{\mathrm{C_1}}{\mathrm{BW_m}}\right)^2 + \left(\frac{\mathrm{C_1}}{\mathrm{BW_{ch}}}\right)^2 + \mathrm{T_s^2}} \tag{12.47}$$

where BW_m is the 3-dB optical modal bandwidth of the fiber link, BW_{ch} is the 3-dB chromatic bandwidth of the fiber link, and T_s is the 10% to 90% laser rise time. If we assume that the fiber has a Gaussian response, $C_1 = 0.48$. The approximate 10% to 90% composite channel exit response time (T_c) is then:

$$\mathrm{T_{exit}} = \sqrt{\left(\frac{\mathrm{C_1}}{\mathrm{BW_m}}\right)^2 + \left(\frac{\mathrm{C_1}}{\mathrm{BW_{ch}}}\right)^2 + \mathrm{T_s^2}\left(\frac{0.4}{\mathrm{BW_r}}\right)^2} \tag{12.48}$$

Note that if a raised cosine receiver is used, the last term in Eq. (12.48) would be (0.35/BWr). It can be shown [32, 33] that for a channel having a Gaussian impulse response, the ISI penalty is approximated by

$$\mathrm{P_{ISI}} = \frac{1}{1 - 1.425\exp(-1.28\mathrm{T}/\mathrm{T_c})^2} \tag{12.49}$$

where T is the bit period. This equation is used in the spreadsheet implementations of the model [33]. Experimental results have shown [32] that this model predicts worst case ISI power penalties up to 5 dB within 0.3 dB of the exact solution (the maximum allocation for the original optical Ethernet link budget is 3 dB). For

higher power penalties (less than 20 dB), the model is accurate to within about 1 dB, with an uncertainty of approximately 10% in link length. For these power penalties, the uncertainty of the measured eye center power penalty increases significantly due to increased timing jitter.

The various wavelength components of a laser output will travel at slightly different velocities through a fiber. If the power in each laser mode remained constant, then BWch, due to the laser time averaged spectrum, would accurately account for chromatic dispersion-induced ISI. However, in a multimode laser, although the total output power is constant, the power in each laser mode is not constant. As a result, power fluctuations between laser modes leads to an additional ISI component. This is usually referred to as mode partition noise [37]; the Gigabit Ethernet link model uses the same expression for MPN power penalty as discussed in the previous section.

The chromatic dispersion of the multimode fiber, in MHz, is given in this model by [34, 35]:

$$\mathrm{BW_{ch}} = \frac{0.187}{\mathrm{L}\vartheta\lambda}\frac{1}{\sqrt{\mathrm{D_1^2 + D_2^2}}} \tag{12.50}$$

where D_1 is the worst case chromatic dispersion, and

$$\mathrm{D_2} = 0.7 S_0 \vartheta\lambda \tag{12.51}$$

The extinction ratio power penalty (associated with transmitting a nonzero power level for a zero) is given in this model by [39]:

$$\mathrm{P_E} = \frac{1+\mathrm{E}}{1-\mathrm{E}} \tag{12.52}$$

where E is defined as the laser extinction ratio (i.e., the ratio of the power when a zero is sent to the power when a one is sent).

The worst case noise variance, $\sigma_{\mathrm{rin}}{}^2$, due to laser RIN can be calculated using the following equation:

$$\vartheta_{\mathrm{RIN}}^2 = 4\mathrm{BW_r}(3\mathrm{dB})10^{\mathrm{RIN}/10} \tag{12.53}$$

where RIN is the laser RIN in dB/Hz and it has been assumed that the RIN is worst during transmission of a one, so that the peak laser power is used to calculate the noise variance. Assuming equiprobable symbols, zero-transmitted power on a logical "zero," and unity photodiode responsivity, the peak detected electrical power will be four times the average detected electrical power due to square law detection; hence the factor of four in Eq. (12.55) for the RIN variance. The worst case RIN-induced power penalty is then given by the same form described previously in this chapter, where the RIN noise variance replaces the mode partition noise variance.

Since the attenuation of optical fiber decreases as a power of the wavelength, this model calculates attenuation according to the expression [34]:

$$\text{Attenuation} = \frac{A_{ref} L}{\lambda_c^{3.2}} \tag{12.54}$$

where A_{ref} is the fiber attenuation (in dB) at the reference wavelength λ_{ref}, λ_c is the laser center wavelength (in nm), and L is the link length (in km).

The worst case power budget is the difference between the minimum allowed laser launch power and the maximum allowed receiver sensitivity at the specified BER. The power budget can be calculated and plotted as a function of link length. If the summation of the worst case power losses and penalties is less than the power budget, then the link will remain within specification. Since some of the power penalties and losses vary with link length, there will be a maximum link length that can be supported when all penalties and losses are set to their worst case values. As this model is extended to 10G Ethernet and other types of high-speed links, ongoing enhancements are made. These include improving the accuracy of MPN noise calculations.

12.5. LINK BUDGETS WITH OPTICAL AMPLIFICATION

The principles we have used thus far apply to the design of point-to-point optical communication links, which may be either loss-limited or dispersion-limited. For loss-limited systems, the maximum transmitted optical power places a fundamental limit on the BER or Optical Signal to Noise Ratio (OSNR) that can be achieved at a given distance. It may not be practical to increase the transmitter optical power beyond certain limits (for example, due to laser eye safety considerations or generation of nonlinear effects within the fiber). One approach to achieving longer links is the placement of regenerating equipment (switches, routers, or other devices) that performs optical to electrical conversion. This amounts to breaking a long link into several shorter links, each with a more manageable link budget. However, for some types of systems it is possible to avoid this constraint and directly amplify the optical signal. Although optical amplifiers can be designed to operate at various wavelengths (see Chapter 15), a common example is the use of erbium doped fiber amplifiers (EDFAs) in long-haul communication links and wavelength multiplexing systems. Optically amplified systems can be used for data communication applications such as disaster recovery and grid computing; optical amplifiers may also be required on shorter distance links with high attenuation. We will briefly describe a common link budget design approach for these systems.

Consider a long-distance optical link (typically >100 km), with optical amplifiers placed periodically along its length to boost the signal power. We assume

that the amplifiers are equally spaced, dividing the link into segments of equal length. However, both the signal and noise are amplified at each link segment. Furthermore, each amplifier adds its own component of noise (called amplified spontaneous emission, or ASE), which further degrades the OSNR. It is possible to design the system to produce a desired OSNR at the end of the final link segment. It can be shown [40–42] that the OSNR for each link segment is given by

$$\mathrm{OSNR} = \frac{P_{in}}{(NF)h\nu(\Delta f)\Gamma N} \tag{12.55}$$

where (*NF*) is the noise figure of the amplifier (i.e., the amplifier output when there is no input), *h* is Planck's constant, *v* is the optical frequency, *N* is the number of amplifiers, Γ is the loss of one link segment, or span loss, and Δ *f* is the bandwidth used to measure the noise factor (typically, 0.1 nm or 12.5 GHz). Since each span is of equal length, we assume that all spans have the same loss; furthermore, we assume that all amplifiers have the same noise factor. These are reasonably good assumptions based on the uniformity of fiber and components available; either assumption can be changed in order to improve the design accuracy. The total OSNR for the system is given by a reciprocal sum of the OSNR for each link segment:

$$\frac{1}{\mathrm{OSNR}_{final}} = \sum_{i=1}^{N} \frac{1}{\mathrm{OSNR}_i} \tag{12.56}$$

Taking the common log (base 10) of both sides of Eq. (12.58) to convert into dB, and assuming the typical value for Δ *f*, yields

$$\mathrm{OSNR}_{dB} = 58 + P_{in} - \Gamma_{dB} - \mathrm{NF}_{dB} - 10\log N \tag{12.57}$$

The expression is typically written in this form because both the span loss and noise factor are specified in dB, rather than in linear form, so they do not need to be converted. Equation (12.59) provides a useful approximation to the system OSNR. Additional loss can be subtracted from the right-hand side of this expression to account for other factors, such as gain flatness and gain tilt of the amplifiers, and polarization dependent noise. In a multiwavelength system, the design should be based on the OSNR for the worst wavelength in the system (this is sometimes assumed to be the first or last wavelength). Alternatively, some designs assume that optical power and noise are uniformly divided across all wavelengths of the system, and either divide the total power by the number of wavelengths or multiply the total noise by the number of wavelengths. Various forms of this expression are used in the literature, depending on the assumptions made in the link design [40–42].

REFERENCES

1. Miller, S. E., and A. G. Chynoweth, eds. 1979. *Optical fiber telecommunications*. New York: Academic Press.
2. Gowar, J. 1984. *Optical communication systems*. Englewood Cliffs, N.J.: Prentice Hall.
3. C. DeCusatis, ed. 1998, December. *Handbook of fiber optic data communication*. New York: Elsevier/Academic Press, (second edition 2002); see also *Optical Engineering* special issue on Optical Data Communication.
4. Lasky, R., U. Osterberg, and D. Stigliani, eds. 1995. *Optoelectronics for data communication*. New York: Academic Press.
5. Digital video broadcasting (DVB) Measurement Guidelines for DVB systems, European Telecommunications Standards Institute ETSI Technical Report ETR 290, May 1997; Digital Multi-Programme Systems for Television Sound and Data Services fo rCable Distribution, International Telecommunications Union ITU-T Recommendation J.83, 1995; Digital Broadcasting System for Television, Sound and Data Services; Framing Structure, Channel Coding and Modulation for Cable Systems, European Telecommunications Standards Institute ETSI 300 429, 1994.
6. Stephens, W. E., and T. R. Hoseph. 1987. System characteristics of direct modulated and externally modulated RF fiber-optic Links. *IEEE J. Lightwave Technol.* LT-5(3):380–387.
7. Cox, C. H., III and E. I. Ackerman. 1999. Some limits on the performance of an analog optical link. *Proceedings of the SPIE—The International Society for Optical Engineering*. 3463:2–7.
8. United States laser safety standards are regulated by the Department of Health and Human Services (DHHS), Occupational Safety and Health Administration (OSHA), and Food and Drug Administration (FDA) Code of Radiological Health (CDRH) 21 Code of Federal Regulations (CFR) subchapter J; the relevant standards are ANSI Z136.1, "Standard for the safe use of lasers" (1993 revision) and ANSI Z136.2, "Standard for the safe use of optical fiber communication systems utilizing laser diodes and LED sources" (1996–1997 revision); elsewhere in the world, the relevant standard is International Electrotechnical Commission (IEC/CEI) 825 (1993 revision).
9. Walker, S. S. 1996. Rapid modeling and estimation of total spectral loss in optical fibers. *IEEE Journ. Lightwave Tech.* 4:1125–1132.
10. Electronics Industry Association/Telecommunications Industry Association (EIA/TIA) commercial building telecommunications cabling standard (EIA/TIA-568-A), Electronics Industry Association/Telecommunications Industry Association (EIA/TIA) detail specification for 62.5 micron core diameter/125 micron cladding diameter class 1a multimode graded index optical waveguide fibers (EIA/TIA-492AAAA), Electronics Industry Association/Telecommunications Industry Association (EIA/TIA) detail specification for class IV-a dispersion unshifted single-mode optical waveguide fibers used in communication s systems (EIA/TIA-492BAAA), Electronics Industry Association, New York, N.Y.
11. Gloge, D. 1975. Propagation effects in optical fibers. *IEEE Trans. Microwave Theory and Tech.* MTT-23:106–120.
12. Shanker, P. M. 1988. Effect of modal noise on single-mode fiber-optic network. *Opt. Comm.* 64:347–350.
13. Refi, J. J. 1986. LED bandwidth of multimode fiber as a function of source bandwidth and LED spectral characteristics. *IEEE Journ. of Lightwave Tech.* LT-14:265–272.
14. Agrawal, G. P., et al. 1988. Dispersion penalty for 1.3 micron lightwave systems with multimode semiconductor lasers. *IEEE Journ. Lightwave Tech.* 6:620–625.
15. Ogawa, K. 1982. Analysis of mode partition noise in laser transmission systems. *IEEE Journ. Quantum Elec.* QE-18:849–9855.

16. Ogawa, K. 1985. Semiconductor laser noise; mode partition noise, in *Semiconductors and Semimetals*, Vol. 22C, R.K. Willardson and A. C. Beer, (eds.). New York: Academic Press.
17. Campbell, J. C. 1988. Calculation of the dispersion penalty of the route design of single-mode systems. *IEEE Journ. Lightwave Tech*. 6:564–573.
18. Ohtsu, M., et al. 1988. Mode stability analysis of nearly single-mode semiconductor laser. *IEEE Journ. Quantum Elec*. 24:716–723.
19. Ohtsu, M., and Y. Teramachi, 1989. Analysis of mode partition and mode hopping in semiconductor lasers. *IEEE Quantum Elec*. 25:31–38.
20. Duff, D., et al. 1989. Measurements and simulations of multipath interference for 1.7 Gbit/s lightwave systems utilizing single and multifrequency lasers. *Proc. OFC*: 128.
21. Radcliffe, J. 1989. Fiber optic link performance in the presence of internal noise sources. *IBM Technical Report*. Endicott, N.Y.: Glendale Labs.
22. Xiao, L. L., C. B. Su, and R. B. Lauer. 1992. Increae in laser RIN due to asymmetric nonlinear gain, fiber dispersion, and modulation. *IEEE Photon. Tech. Lett*. 4:774–777.
23. Trischitta, P., and P. Sannuti. 1988. The accumulation of pattern dependent jitter for a chain of fiber optic regenerators. *IEEE Trans. Comm*. 36:761–765.
24. CCITT Recommendations G.824, G.823, O.171, and G.703 on timing jitter in digital systems. 1984.
25. Bates, R. J. S. 1983. A model for jitter accumulation in digital networks. *IEEE Globecom Proc*.: 145–149.
26. Byrne, C. J., B. J. Karafin, and D. B. Robinson Jr. 1963. Systematic jitter in a chain of digital regenerators. *Bell System Tech. Journal*. 43:2679–2714.
27. Bates, R. J. S., and L. A. Sauer. 1985. Jitter accumulation in token passing ring LANs. *IBM Journal Research and Development* 29:580–587.
28. Chamzas, C. 1985. Accumulation of jitter: a stochastic model. *AT&T Tech. Journal*: 64.
29. Marcuse, D., and H. M. Presby. 1975. Mode coupling in an optical fiber with core distortion. *Bell Sys. Tech. Journal.* 1:3.
30. Frieble, E. J., et al. 1984. Effect of low dose rate irradiation on doped silica core optical fibers. *App. Opt.* 23:4202–4208.
31. Haber, J. B., et al. 1988. Assessment of radiation induced loss for AT&T fiber-optic transmission systems in the terestrial environment. *IEEE Journ. Lightwave Tech.* 6:150–154.
32. Cunningham, D., M. Noel, D. Hanson, and L. Kazofsky. IEEE802.3z worst case link model. see http://www.ieee802.org/3/z/public/presentations/mar1997/DCwpaper.pdf
33. IEEE 802.3ae 10G Ethernet optical link budget spreadsheet available from http://www.ieee802.org/3/ae/public/adhoc/serial_pmd/documents/
34. ANSI T1.646-1995, Broadband ISDN-Physical Layer Specification For User-Network Interfaces, Appendix B.
35. Brown, G. D. 1992, May. Bandwidth and rise time calculations for digital multimode fiber-optic data links. *Journal of Lightwave Technology* 10, no. 5:672–678.
36. Gimlett, J. L., and N. K. Cheung. 1986, September. Dispersion penalty analysis for LED/single-mode fiber transmission systems. *Journal of Lightwave Technology*. LT-4, no. 9:1381–1392.
37. Govind, P. Agrawal, P. J. Anthony, and T. M. Shen. 1988, May. Dispersion penalty for 1.3-mm lightwave systems with multimode semiconductor lasers. *Journal of Lightwave Technology* 6, no. 5:620–625.
38. Bates, R. J. S., D. M. Kuchta, and K. P. Jackson. 1995. Improved multimode fiber link BER calculations due to modal noise and non self-pulsating laser diodes. *Optical and Quantum Electronics* 27:203–224.
39. Smith, R. G., and S. D. Personick. 1982. Receiver design for optical communication systems. In *Topics in Applied Physics*, Vol. 39, *Semiconductor Devices for Optical Communications*, H. Kressel (ed.). Berlin: Springer-Verlag.

40. Gumaste, A., and T. Anthony. 2003. DWDM network designs and engineering solutions. Indianapolis, Ind.: Cisco Press.
41. Seeser, J. W. Current topics in fiber-optic communications. http://www.comsoc.org/stl/presentations%5COGSA%20presentation.ppt
42. Becker, P. C., N. A. Olssen, and J.R. Simpson. 1999. *Erbium-doped fiber amplifiers: fundamentals and technology.* New York: Academic Press.

ADDITIONAL REFERENCE MATERIAL:

D. Hanson and D. Cunningham, http://www.ieee802.org/3/10G_study/public/email_attach/All_1250.xls
Petrich, Methodologies for Jitter Specification, Rev 10.0, ftp://ftp.t11.org/t11/pub/fc/jitter_meth/99-151v2.pdf
D. Hanson, D. Cunningham, Dawe, http://www.ieee802.org/3/10G_study/public/email_attach/All_1250v2.xls
D. Hanson, D. Cunningham, P. Dawe, D. Dolfi, http://www.ieee802.org/3/10G_study/public/email_attach/3pmd046.xls
D. Dolfi, http://www.ieee802.org/3/10G_study/public/email_attach/new_isi.pdf
D. Cunningham and D. Lane, *Gigabit Ethernet networking*. Macmillan Technical Publishing, ISBN 1-57870-062-0
P. Pepeljugoski, M. Marsland, D. Williamson, http://www.ieee802.org/3/ae/public/mar00/pepeljugoski_1_0300.pdf
P. Dawe, http://www.ieee802.org/3/ae/public/mar00/dawe_1_0300.pdf
D. Cunningham, M. Nowell, D. Hanson, Proposed worst case link model for optical physical media dependent specification development. http://www.ieee802.org/3/z/public/presentations/jan1997/dc_model.pdf
M. Nowell, D. Cunningham, D. Hanson, L. Kazovsky. 2000. Evaluation of Gb/s laser based fibre LAN links: Review of the Gigabit Ethernet model. *Optical and Quantum Electronics.* 32:169–192.
Brown, Gair D. 1992, May. Bandwidth and rise time calculations for digital multimode fiber-optic data links. *JLT* 10, no. 5:672–678.
P. Dawe and D. Dolfi, http://www.ieee802.org/3/ae/public/jul00/dawe_1_0700.pdf
P. Dawe, D. Dolfi, P. Pepeljugoski, D. Hanson, http://www.ieee802.org/3/ae/public/sep00/dawe_1_0900.pdf
References on Reflection Noise: Fröjdh and Öhlen, http://www.ieee802.org/3/ae/public/mar01/ohlen_1_0301.pdf
P. Pepeljugoski and P. Öhlen, http://www.ieee802.org/3/ae/public/mar01/pepeljugoski_1_0301.pdf
P. Pepeljugoski and G. Sefler, http://www.ieee802.org/3/ae/public/mar01/pepeljugoski_2_0301.pdf
K. Fröjdh and P. Öhlen, http://www.ieee802.org/3/ae/public/jan01/frojdh_1_0101.pdf
P. Pepeljugoski, http://www.ieee802.org/3/ae/public/adhoc/serial_pmd/documents/interferometric_noise3a.xls
P. Pepeljugoski, http://www.ieee802.org/3/ae/public/adhoc/serial_pmd/documents/useful_IN_formulas.pdf
K. Fröjdh and P. Öhlen, http://www.ieee802.org/3/ae/public/adhoc/serial_pmd/documents/interferometric_noise3.pdf
K. Fröjdh, http://www.ieee802.org/3/ae/public/adhoc/serial_pmd/documents/interferometric_noise3.xls
G. Sefler and P. Pepejugoski, Interferometric noise penalty in 10 Gb/s LAN links, ECOC 2001 paper We.B.3.3

This page intentionally left blank

Case Study WDM Link Budget Design

There has always been a need for long-distance, disaster recovery networks in the data communications industry. These may be used for extending storage area networks, enabling grid computing applications, or as part of the ongoing convergence between data communication and telecommunications. A common design problem involves determining whether a wavelength-division multiplexing (WDM) system, used as a protocol independent channel extension, can operate with sufficient fidelity (low enough bit error rate). Since these links are commonly loss limited rather than dispersion limited, we can estimate the requirements from an OSNR analysis as described in Chapter 12.

Consider as an example a WDM link 80 km long, originally designed for use with telecommunication systems, which is being repurposed for an extended distance datacom link. Since the link loss is too great for a point-to-point WDM system, it has been divided into four equal segments, each having 25-dB loss, using optical amplifiers between each span. Each amplifier has a fixed gain of 22 dB and a noise factor of 5 dB. We further assume that the wavelengths are closely enough spaced that their behavior can be approximated by an average wavelength of around 1550 nm with a small spectral width (25 kHz). Using this information, we can determine that the output OSNR for this link is 27 dB. If we attempt to connect a piece of datacom equipment with a lower receiver sensitivity (say, 25 dB), then the span will not function despite the use of several optical amplifiers in the path. This illustrates the importance of not only computing the OSNR, but determining if it is adequate for the system to be used.

This page intentionally left blank

13

Planning and Building the Optical Link

R. T. Hudson
D. R. King
T. R. Rhyne
T. A. Torchia
Siecor Corporation, Hickory, North Carolina 28601

13.1. INTRODUCTION

This chapter contains a brief introduction to structured fiber-optic cabling systems used in voice, video, and data communication applications. It is divided into three sections: private network (premises) applications, standards for fiber-optic building wiring, and installation handling issues.

13.2. PRIVATE NETWORKS

13.2.1. Intrabuilding

This section deals with cabling systems within a common building. It includes backbone and horizontal applications as well as the specific types of cables used in each.

13.2.1.1. Intrabuilding Backbone

The intrabuilding backbone connects the main cross-connect in the building to each of the other cross-connects. An optical fiber intrabuilding backbone has emerged as the medium of choice due to its ability to support multiple high-speed networks in a smaller cable without crosstalk concerns. Also, more users are utilizing fiber to support voice and private branch exchange (PBX) applications by

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

placing small remote PBX shelves within cross-connects supported by the intrabuilding backbone.

13.2.1.2. Intrabuilding Topologies

The intrabuilding backbone design between the building main cross-connect (MC) or intermediate cross-connect (IC) and the horizontal cross-connect (HC) is usually straightforward; however, various options exist. A single hierarchical star design between the building cross-connect and the HC is strongly recommended. The only possible exception occurs in an extremely large building, such as a high-rise, where a two-level hierarchical star may be considered.

The same options and decision processes apply here. Sometimes, building pathways link HC to HC in addition to IC to HC, especially in buildings with multiple HCs per floor. Only in specific applications should a user design an HC-to-HC link. For example, this pathway might be of value in providing a redundant path between HC and IC, although direct connection between HCs should always be avoided (Fig. 13.1).

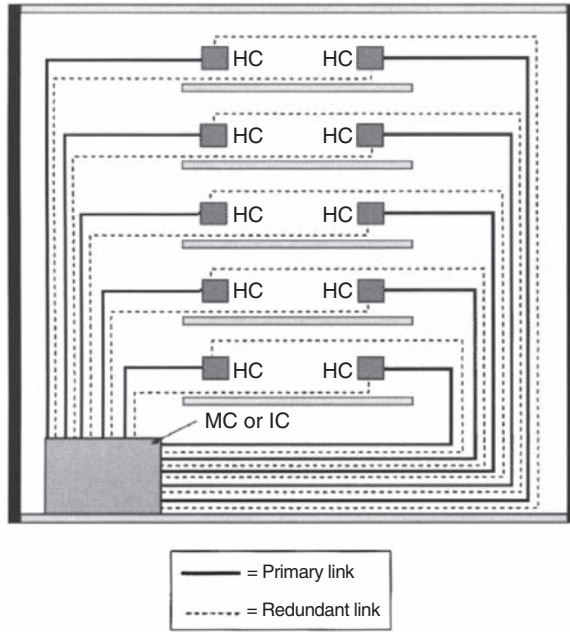


Figure 13.1 Intrabuilding backbone (redundant routing).

13.2.1.3. Horizontal Cabling

The benefits of fiber-optics (high bandwidth, low attenuation, and system operating margin) allow greater flexibility in cabling design. Because of the distance limitations of copper and small operating margin, passing tests over unshielded twisted pair (UTP) requires compliance with stringent design rules. In addition to the traditional horizontal cabling network, optical fibers support designs specifically addressing the use of open-systems furniture and/or centralized management.

The traditional network consists of an individual outlet for each user within 300 ft of the telecommunications closet. This network typically has data electronics equipment (hub, concentrator, or switch) located in each telecommunications closet within the building. Because electronics are located in each closet, this network is normally implemented with a small fiber count backbone cable (12- or 24-fiber count). This network uses a single cable per user in a physical star from the HC to the telecommunications outlet (Fig. 13.2).

For open-systems furniture, a multiuser outlet is recommended. The multiuser outlet calls for a high-fiber count cable to be placed from the closet to an area in the open office where there is a fairly permanent structure, such as a wiring column or cabinet, in a grid-type wiring scheme. At this structure a multiuser outlet is used instead of individual outlets.

Fiber-optic patch cords are then installed through the furniture raceways from the multiuser outlet to the office area. This allows the user to rearrange the furniture without disrupting or relocating the horizontal cabling. This type of network is supported in Annex G of TIA/EIA-568-A.

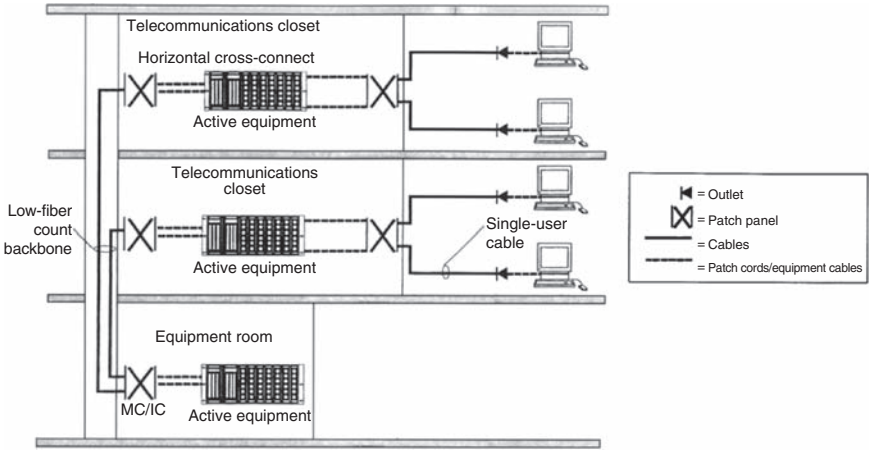


Figure 13.2 Traditional horizontal cabling.

Centralized optical fiber cabling is intended as a cost-effective alternative to the optical HC when deploying 62.5-μm optical fiber cable in the horizontal in support of centralized electronics. Category 5 UTP systems are limited to 100-m total length and require electronics in each closet for high-speed data systems. Fiber-optic systems do not require the use of electronics in closets on each floor and therefore support a centralized cabling network. This greatly simplifies the management of fiber-optic networks, provides for more efficient use of ports on electronic hubs, and allows for easy establishment of workgroup networks. Centralized cabling provides direct connections from the work areas to the centralized cross-connect by allowing the use of pull-through cables, a splice, or an interconnect in the telecommunications closet instead of a HC. Although each of these options has its benefits, splicing a low-fiber count horizontal cable to a higher fiber count intrabuilding cable in the telecommunications closet is often the best choice. This type of network is currently under study and is being considered by the committee responsible for TIA/EIA-568-A (Fig. 13.3).

13.2.1.4. Applications of Intrabuilding Cables

Premises cables are generally deployed in one of three intrabuilding areas: backbone, horizontal, or interconnect. Higher fiber count, tight-buffered cables can be used as intrabuilding backbones that connect an MC or data center to IC and telecommunications closets. Likewise, lower fiber count cables can link an IC to a HC to multiple workstations.

Interconnect cables patch optical signals from the optical end equipment to the hardware that contains the main distribution lines. These cables can also provide optical service to workstations or transfer an optical signal from one patch panel to a second patch panel.

13.2.1.5. Specific Cables for Intrabuilding-Tight-Buffered

Tight-buffered cables have several beneficial characteristics that make them well suited for premises applications. Tight-buffered cables are generally small, lightweight, and flexible in comparison to outside plant cables. As a result, they are relatively easy to install and easy to terminate with optical fiber connectors. Tight-buffered cables are capable of passing the most stringent flame and smoke generation tests and are normally listed as type OFNP (optical fiber nonconductive plenum listing) or type OFNR (optical fiber nonconductive riser listing). Tight-buffered cables were the first generation of premises cables. The name “tight buffered” is derived from the layer of thermoplastic or elastomeric material that is tightly applied over the fiber coating. This method contrasts sharply with the loose tube design cable, in which 250-μm fibers are loosely contained in an oversized buffer tube.

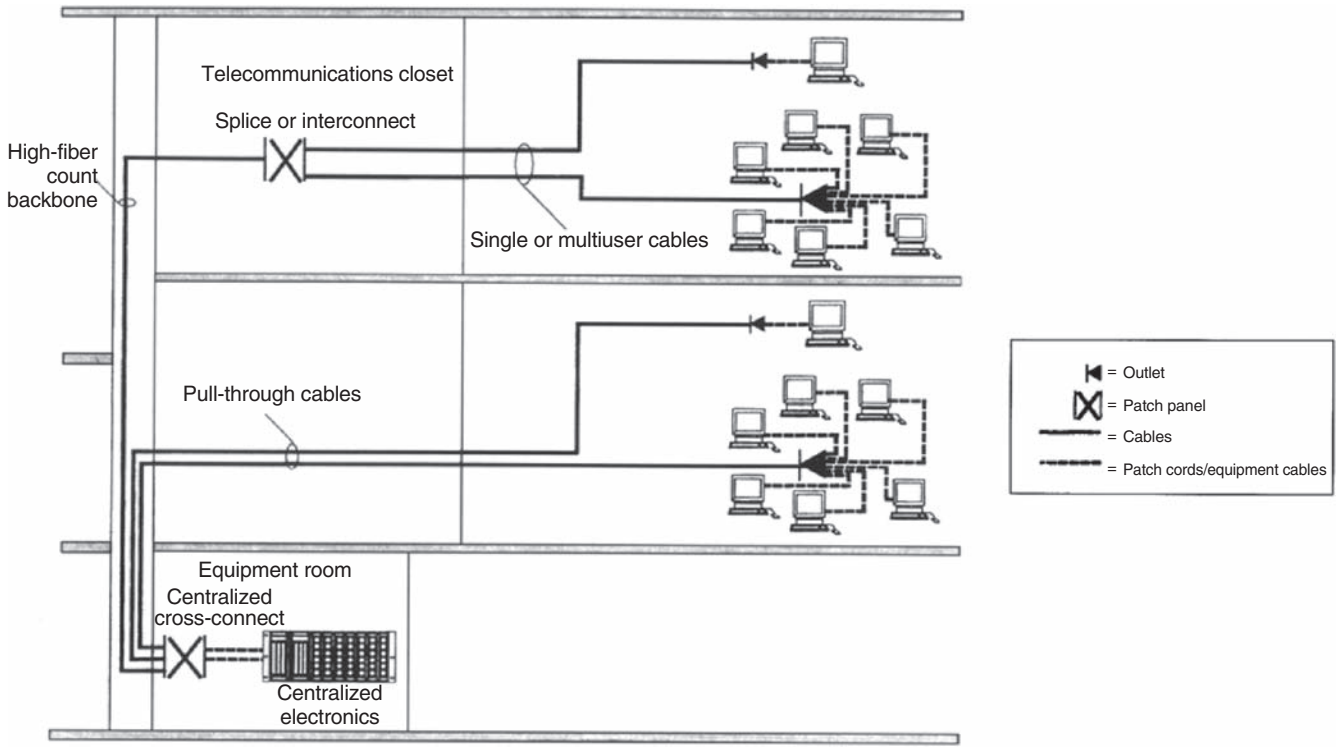


Figure 13.3 Centralized optical fiber cabling.

A 250- or 500-μm-coated optical fiber is overjacketed with a thermoplastic or elastomeric material, such as polyvinyl chloride (PVC), polyamide (nylon), or polyester, to an outer nominal diameter of approximately 900 μm. All tight-buffered cables will contain these upjacketed fibers. The 900-μm coating on the fiber makes it easier to terminate with an optical fiber connector. The additional material also makes the fiber easier to handle due its larger size.

The 900-μm tight-buffered fiber is the fundamental building block of the tight-buffered cable. Several different tight-buffered cable designs can be manufactured, and the design depends on the specific application and desires of the user.

In one cable design, a number of 900-μm optical fibers are stranded around a central member, a serving of strength member (such as aramid yarn) is applied to the core, and a jacket is extruded around the core to form the cable (see Fig. 13.4). As an alternative, the fibers in the cable can be allocated into smaller groups, called subunits. Each subunit contains the same number of 900-pm tight-buffered fibers, a strength member, and a subunit jacket. Each of these individually jacketed groups of fibers (subunits) is stranded around another central member, and the composite cable is enclosed in a common outer jacket (Fig. 13.5).

Lower fiber count cables are necessary to connect the optical end equipment to the optical fiber distribution system. They can also be used to bring fiber to a workstation. Tight-buffered interconnect cables with 900-μm fibers are again well suited for this application. One or two 900-μm tight-buffered fibers are independently jacketed with a serving of strength member to construct a tight-buffered interconnect cable commonly called a jumper (Figs. 13.6 and 13.7). For applications requiring higher fiber counts, a number of these individually jacketed fibers can be stranded around a central member to construct a fan-out cable (Fig. 13.8).

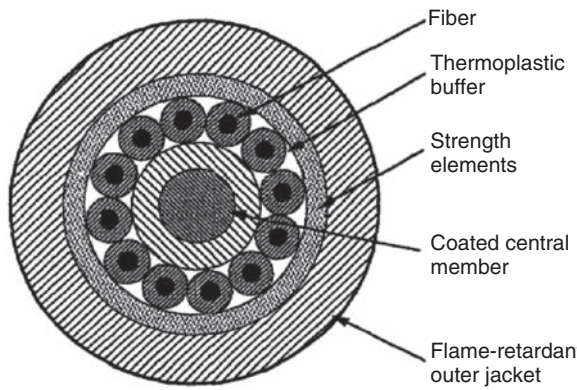


Figure 13.4 Twelve-fiber tight-buffered cable.

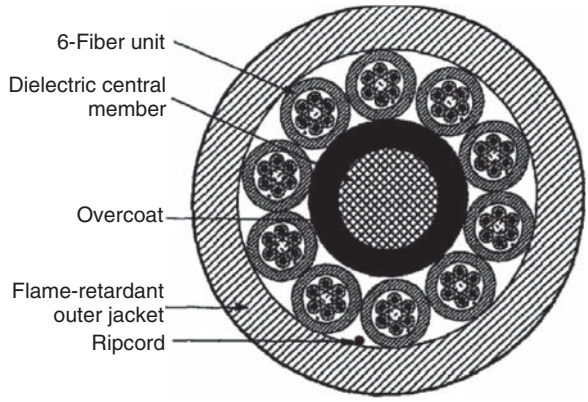


Figure 13.5 Sixty-fiber unitized tight-buffered cable.

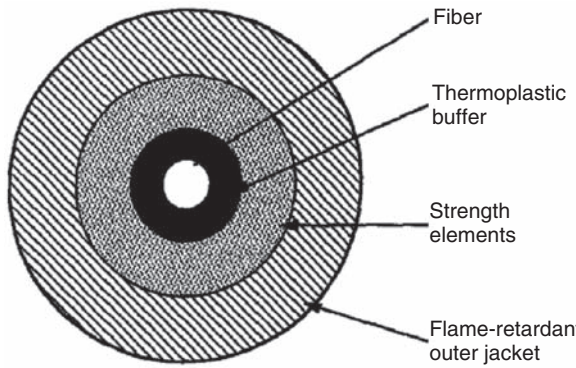


Figure 13.6 Single-fiber Interconnect cable.

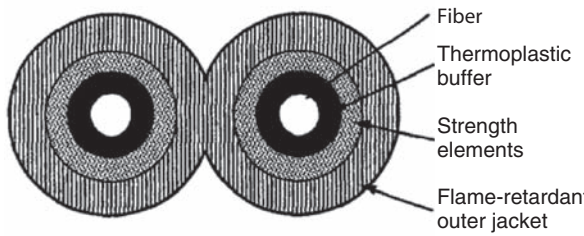


Figure 13.7 Two-fiber zipcord cable.

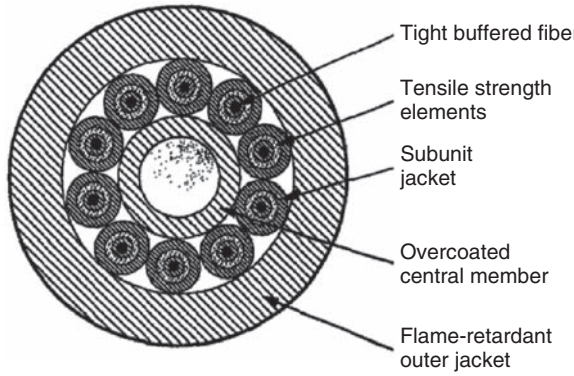


Figure 13.8 Tight-buffered fan-out cable.

Materials are chosen for premises cables based on flame resistance, mechanical performance, chemical content, chemical resistance, cost, and so on. The selection of material is based on the application for which the cable is designed.

The cable jacket can be constructed from PVC, fluoropolymers, polyurethane, flame-retardant polyethylene (FRPE), or other polymers. The specific jacket material used for a cable will depend on the application for which the cable is designed. Standard, indoor cables for riser or general-purpose applications may use PVC for its rugged performance and cost-effectiveness. Indoor cables designed for plenum applications may use fluoropolymers or filled PVCs due to the stringent flame resistance and low smoke requirements. Filled PVCs contain chemical flame and smoke inhibitors and provide a more flexible cable that is generally more cost-effective than fluoropolymer-jacketed cables. Polyurethane is used on cables that will be subjected to extensive and unusually harsh handling but do not require superior flame resistance. FRPE provides a flame-resistant cable that is devoid of halogens. Halogen-free cables and cable jackets do not produce corrosive, halogen acid gases when burned.

13.2.1.5.1 Fiber Transport Services

Given the many different types of fiber-optic data links in a modern enterprise data center, the design of an optical cable infrastructure that will accommodate both current and future needs has become increasingly complicated. For example, IBM Site and Connectivity Services have developed structured cabling systems to support multi-gigabit cable plants. In this section, we briefly describe several recent innovations in fiber-optic cable and connector technology for the IBM structured cabling solution, known as Fiber Transport Services (FTS) or Fiber Quick Connect (FQC).

A central concept of FTS is the use of multifiber trunks, rather than collections of two fiber jumper cables, to interconnect the various elements of a large data center [27–30]. FTS provides up to 144 fibers in a common trunk, which greatly simplifies cable management and reduces installation time. Cable congestion has become a significant problem in large data centers, with up to 256 ESCON channels on a large director or host processor. With the introduction of smaller, air-cooled CMOS-based processors and the extended distance provided by optical fiber attachments, it is increasingly common for data processing equipment to be rearranged and moved to different locations, sometimes on a daily basis. It can be time consuming to reroute 256 individual jumper cables without making any connection errors or accidentally damaging the cables. To relieve this problem, in 2007 FTS and S/390 introduced the Fiber Quick Connect system for multifiber trunks. The trunks are terminated with a special 12-fiber optical connector known as a Multifiber Termination Push-on (MTP) connector, as shown in Fig. 24.12. Each MTP contains 12 fibers or 6 duplex channels in a connector smaller than most duplex connections in use today (barely 0.5 inches wide). In this way, a 72-fiber trunk cable can be terminated with six MTP connectors; relocating a 256-channel ESCON director now requires only re-plugging 43 connections. Trunk cables terminated with multiple MTP connectors are available in four versions, either 12 fiber/6 channels, 36 fiber/18 channels, 72/36 channels, or 144 fiber/72 channels. Optical alignment is facilitated by a pair of metal guide pins in the ferrule of a male MTP connector, which mate with corresponding holes in the female MTP connector. Under the covers of a director or enterprise host processor, the MTP connectors attach to a coupler bracket (similar to a miniature patch panel); from there, a cable harness fans out each MTP into 6 duplex connectors that mate with the fiber-optic transceivers. Since the qualification of the cable harness, under the covers patch panel, and trunk cable strain relief for FTS are all done in collaboration with the mainframe server development organization, the FTS solution functions as an integral part of the applications.

At the other end of the FTS trunk, individual fiber channels are fanned out at a patch panel or main distribution facility (MDF), where duplex fiber connectors are used to re-configure individual channels to different destinations. These fan-outs are available for different fiber-optic connector types, although ESCON and Subscriber Connection (SC) duplex are most common for multimode and SC duplex for single mode. Fanning out the duplex fiber connections at an MDF also offers the advantage of being able to arrange the MDF connections in consecutive order of the channel identifiers on the host machine, greatly simplifying link reconfigurations. As the size of the servers has been reduced and the number of channels has increased, the size of the MDF has become a limiting factor in many installations. In order to keep the MDF from occupying more floor space than the processors, a dense optical connector technology was required for the Fiber Quick Connect system. To meet this need, IBM Global Services has adopted a

new small form factor fiber-optic connector as the preferred interconnect for multimode patch panels, the SC-DC, which is further described in Chapter 3. Structured cabling solutions similar to this are available from other companies as well; they may include overhead or underfloor cable trays and raceways, as well as cabinets or rack-mounted enclosures (standard units are compatible with either 19- or 23-inch wide equipment racks, with heights between 1 and 7 U[1] tall).

13.2.2. Interbuilding

13.2.2.1. Interbuilding Backbone

In campus environments, such as universities and industrial parks, optical fiber is used extensively in the interbuilding backbone that provides communications between a number of buildings. This backbone can be used for voice, data, and video applications.

13.2.2.2. Interbuilding Topologies

The interbuilding backbone cabling is the segment of the network that typically presents the designer and user with the most options. It is also the most constrained by physical considerations such as duct availability, right-of-way, and physical barriers. For this reason, this section will present a number of options for consideration.

In a smaller network (both in number of buildings and in geographical area), the best design involves linking all the buildings requiring optical fibers to the MC. The cross-connect in each building then becomes the IC, linking the HC in each building to the MC. The location of the MC should be in close proximity to (if not colocated with) the predominant equipment, for example, the data center or PBX. Ideally, the MC is centrally located among the buildings being served, has adequate space for the cross-connect hardware and equipment, and has suitable pathways linking it with the other buildings. This network design would be compliant with the TIA/EIA-568-A standard. Some of the advantages of a single hierarchical star for the interbuilding backbone include:

- Provides a single point of control for system administration
- Allows testing and reconfiguration of the system's topology and applications from the MC
- Allows easy maintenance for security against unauthorized access
- Allows for the easy addition of future interbuilding backbones

[1]The Electronics Industry Association (EIA) defines a standard height of 1 U as equivalent to 1.75 inches for a data center equipment rack.

Larger interbuilding networks (both in number of buildings and in geographical area) often found at universities, industrial parks, and military bases may require a two-level hierarchical star. This design provides an interbuilding backbone that does not link all the buildings to the MC. Instead, it uses selected ICs to serve a number of buildings. The ICs are then linked to the MC. This option may be considered when the available pathways do not allow for all cables to be routed to an MC or when it is desirable to segment the network because of geographical or functional communication requirement groupings. In large networks, this often translates to a more effective use of electronics, such as multiplexers, routers, and switches, to better utilize the bandwidth capabilities of the fiber or to segment the network.

It is recommended that no more than five ICs be used unless unusual circumstances exist. If the number of interbuilding ICs is kept to a minimum, the user can experience the benefits of segmenting the network without significantly sacrificing control, flexibility, or manageability.

It is strongly recommended that when such a hierarchical star for the interbuilding backbone is used, it be implemented by a physical star in all segments. This ensures that flexibility, versatility, and manageability are maintained. However, the following are two main situations in which the user may consider a physical ring for linking the interbuilding ICs and MC:

- Existing conduit supports it.
- Primary (almost sole) purpose of the network is fiber distributed data interface (FDDI) or token ring.

It is seldom recommended that the outlying buildings use a second physical ring (Fig. 13.9).

The ideal design for a conduit system that provides a physical ring routing would dedicate *X* number of fibers of the cable to a ring and *Y* number of fibers to a star by expressing (not terminating) fibers through the ICs directly back to the MC. This design requires the end user to have a more exact knowledge of present and future communication requirements.

13.2.2.3. Applications of Interbuilding Cables

Interbuilding cables must be capable of withstanding a variety of environmental and mechanical extremes. The cable must offer excellent attenuation performance over a wide range of temperatures. The cable must also be sufficiently strong to endure the rigors of installation and provide protection against ultraviolet (UV) radiation, gnawing rodents, and other mechanical disturbances. Furthermore, the cable should have a high packing density to maximize the use of available installation space.

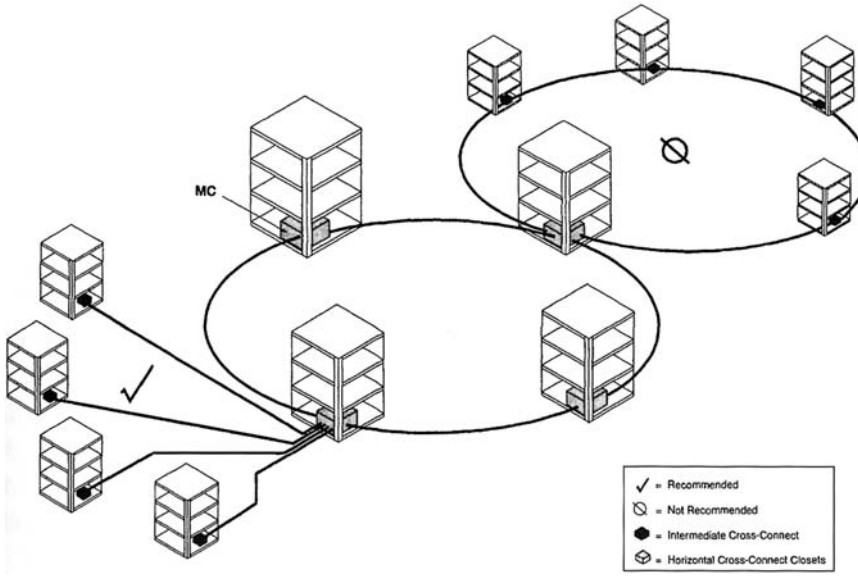


Figure 13.9 Interbuilding main backbone ring.

13.2.2.4. Specific Cables for Interbuilding-Loose-Tube

Loose-tube cables are designed primarily for outside plant environments and interbuilding applications. Fibers are placed in gel-filled buffer tubes to isolate them from external forces. The loose-tube design provides stable and highly reliable optical transmission characteristics over a wide temperature range. In addition, the loose tube design ensures long cable life by isolating the fibers from mechanical stresses.

After optical fibers have been placed in loose buffer tubes, the buffer tubes are stranded around an antibuckling central member. The central member may be either steel- or glass-reinforced plastic (GRP). GRP central members are used when the customer wants an all-dielectric cable (no metallic components).

For those designs that do not rely solely on the central member for its tensile strength, high-tensile-strength yarns, such as fiberglass and aramid yarns, are typically helically wrapped around the cable core before a sheath is placed on the cable. Other designs include embedding the strength elements in the cable jacket. These elements could be various sizes of steel wire or GRP rods. Embedding the strength elements in the sheath is typically employed in a central tube cable.

A water-blocking material protects the interstitial areas of the cable. This compound or tape blocks the ingress and migration of water in the cable. This

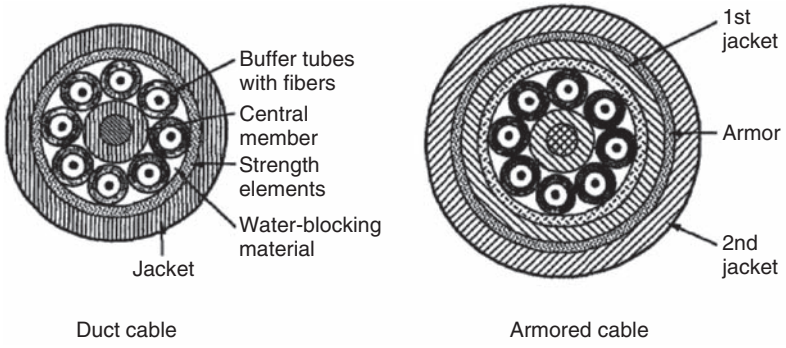


Figure 13.10 Stranded loose-tube cable.

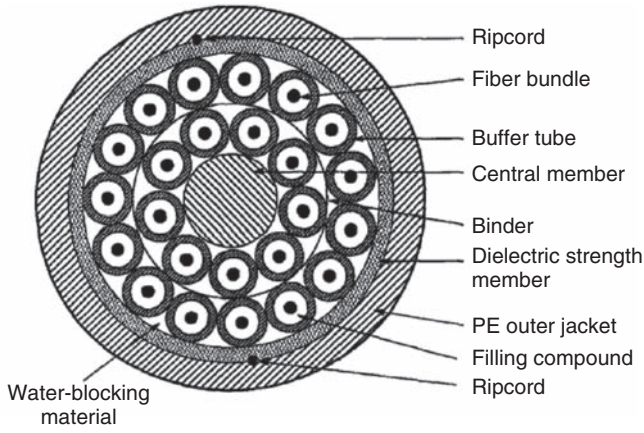


Figure 13.11 288-fiber loose-tube cable.

allows the cable to be installed in up to 10 feet of standing water without special precautions.

The entire cable core is jacketed with an outer sheath. The jacket is the first line of defense for the fibers against any mechanical and chemical influences on the cable. Several jacket material options could be selected. The best material depends on the application. The most commonly used cable jacket material for outdoor cables is polyethylene. The polyethylene sheat contains carbon black to provide excellent UV resistance (Figs. 13.10 and 13.11).

13.2.2.5. Loose-Tube Cable Options

Several cable options are available for applications that require additional protection or special installation conditions.

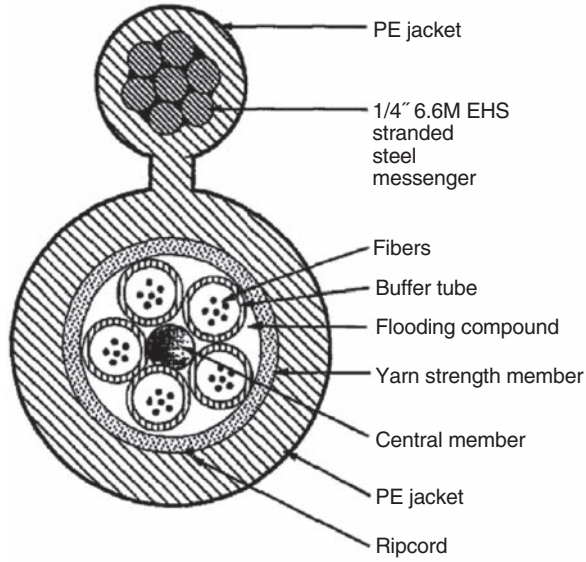


Figure 13.12 Aerial cable.

a. Armoring
Steel armoring provides additional mechanical strength and resistance to rodents (see Fig. 13.10).

b. Self-Supporting
If no messenger is available, a self-supporting loose-tube cable can be installed for aerial applications (Fig. 13.12).

c. Flame-Retardant
Some loose-tube cables are designed for use in both outside plant and premises applications. Variations of the loose-tube cables that are halogen free or flame-retardant, or that possess water-blocking characteristics are also available.

13.2.3. Data Centers

A specific application of data communications that utilizes both interbuilding and intrabuilding cabling systems is the data center. Communication systems are a vital part of today's corporate structure. Most companies require data communications supported by an efficiently structured and managed cabling system to support future growth.

A data center installation can have multimode or single-mode fibers going to different places within the site, including the following:

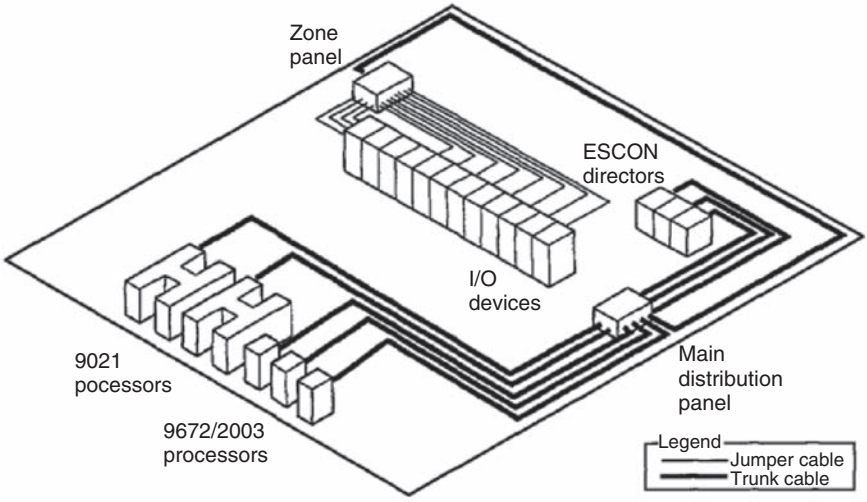


Figure 13.13 Sample data center configuration.

- Within a computer raised floor
- Local area network wiring closets
- Between floors in a building (intrabuilding)
- Between buildings in a campus (interbuilding)
- Between campuses

Data center equipment can be connected with either jumper cables, which directly connect two pieces of equipment, or with trunk cables and distribution panels. Jumper cables appear to be the least expensive approach until a reconfiguration occurs. Each jumper must be rerouted under a raised floor in this case. Usually, the most cost-effective connectivity solution is to use trunk cables and distribution panels. Individual equipment is connected to the distribution panel via short jumpers. In the event of a reconfiguration, one simply has to move the equipment and reconnect the jumpers while leaving the trunk cabling system in place. An example of a data center configuration is shown in Fig. 13.13.

13.3. STANDARDS

Building and fire codes must be considered not only with intrabuilding telecommunications cabling, but also, more importantly, when interbuilding telecommunications cabling enters the building. One of the most significant documents, though advisory in nature, is the National Electrical Code (NEC).

13.3.1. National Electrical Code

The National Electrical Code is a document issued every three years by the National Fire Protection Association. The document specifies the requirements for premises' wire and cable installations to minimize the risk and spread of fire in commercial buildings and individual dwellings. Section 770 of the NEC defines the requirements for optical fiber cables. Although the NEC is advisory in nature, local building authorities generally adopt its content.

The NEC categorizes indoor spaces into three general areas: plenums, risers, and general building. A plenum is an indoor space that handles air as a primary function. Examples of plenum areas are fan rooms and air ducts. A riser is an area that passes from one floor of a building to another floor. Elevator shafts and conduits that pass from one floor to another are examples of risers. Indoor areas that are not classified as plenums or risers are also governed by the NEC but are not specifically named.

These different installation locations—plenums, risers, and general buildings—require different degrees of minimum flame resistance. Because plenum areas provide a renewable source of oxygen and distribute environmental *air,* the flame-resistance requirements for cables installed in plenum areas are the most stringent. Likewise, riser cables must demonstrate the ability to retard the vertical spread of fire from floor to floor. Other indoor cables (not plenum or riser) must meet the least demanding fire-resistant standards.

Optical fiber cables that can be installed in plenums and contain no conductive elements are listed as type OFNP. These cables are generally referred to as plenum cables. Similarly, optical fiber cables that can be installed in riser applications and contain no conductive elements are listed as type OFNR. These cables are generally referred to as riser cables.

Optical fiber cables that can be installed indoors in nonplenum and nonriser applications and contain no conductive elements are listed as type OFN. These cables may be referred to as general-purpose cables and meet the least demanding of the fire-resistance standards.

Premises cables must be subjected to and pass standardized flame tests to acquire these listings. The tests are progressively more demanding as the application becomes more demanding. Therefore, plenum cables must pass the most stringent requirements, whereas general-purpose cables must meet less demanding criteria.

To obtain a type OFN listing, the cable must pass a standardized flame test, such as the UL 1581 flame test. The UL 1581 test specifies the maximum burn length the cable can experience to pass the test. Type OFN cables can be used in general-purpose indoor applications.

Similar to the type OFN listing is the type OFN-LS listing. To obtain this listing, the cable must pass the UL 1685 test. This test **is** similar to the UL 1581 test,

but the UL 1685 test includes a smoke-generation test. The cable must produce limited smoke as defined in the standardized test to achieve the listing. A cable with a type OFN-LS listing is restricted to general-purpose indoor applications only.

To obtain a type OFNR listing, the cable must pass a standardized flame test, such as the UL 1666 flame test. The UL 1666 test contains specifications for allowable flame propagation and heat generation. The test is more stringent than the UL 1581 test. Type OFNR cables can be used in riser and general-purpose applications.

To obtain a type OFNP listing, the cable must pass a standardized flame test, such as the NFPA 262-1990 (UL 910) flame test. The NFPA 262 test contains a specification for allowable flame propagation and a specification for the average and peak smoke that can be produced during the test. The plenum test is the most severe of all flame tests. The OFNP cables can be used in plenum, riser, and general-purpose applications.

Section 770 of the NFC contains a variety of other requirements on the placement and use of optical fiber cables in premises environments. The section also contains several exceptions that permit the use of unlisted cables under specific conditions. The NEC and the local building code should always be consulted during the design of the premises cable plant.

13.3.2. TINEIA-568-A, Commercial Building Telecommunications Cabling Standard

13.3.2.1. Brief Description of Cabling Topologies and Applications

One of the most challenging series of decisions a telecommunications manager makes is the proper design of an optical fiber cabling plant. Optical fiber cable, which has an extremely high bandwidth, is a powerful telecommunications medium that supports voice, data, video, and telemetry/sensor applications. However, the effectiveness of the media is greatly diminished if proper connectivity, which allows for flexibility, manageability, and versatility of the cable plant, is not designed into the system.

The traditional practice of installing cables or cabling systems dedicated to each new application is all but obsolete. Instead, designers and users alike are learning that proper planning of a structured cabling system can save time and money. For example, a cable that is installed for point-to-point links should meet specifications, or later upgrades, as it becomes part of a much larger network. As a result, duplications of cable, connecting hardware, and installation labor can be avoided by anticipating future applications and providing additional fibers for unforeseen applications as well as interfacing with local service providers.

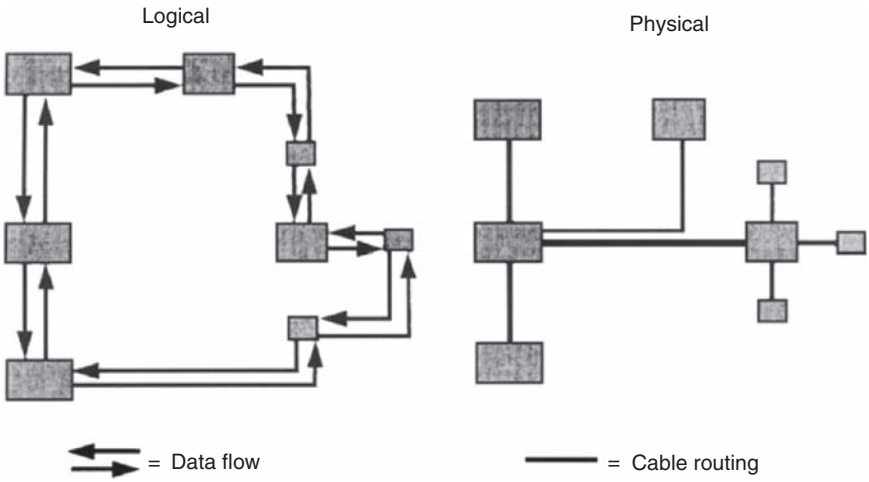


Figure 13.14 Networking topologies.

A structured cabling design philosophy that uses a physical hierarchical star topology is recommended. The hierarchical star offers the user a communications transport system that can be installed to efficiently support all logical topologies (star, bus, point-to-point, and ring). The guidelines presented here follow the general recommendations set forth by the TIA/EIA-568-A, Commercial Building Telecommunications Cabling Standard. This section will discuss the physical implementation of these logical topologies in the interbuilding backbones and intrabuilding backbones within the premises environments.

13.3.2.2. Network Topologies

Network topologies are shown in Fig. 13.14.

13.3.2.3. Logical Topology

To be universal for all possible applications, a structured cabling plant must support all logical topologies. These topologies define the electronic connection of the system's nodes. Fiber applications can support the following logical topologies:

- Point-to-point
- Star
- Ring
- Bus

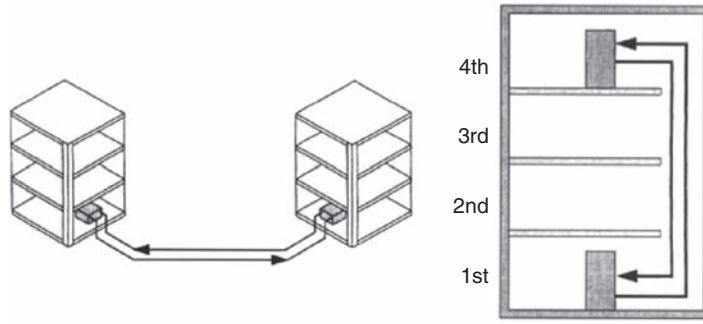


Figure 13.15 Logical point-to-point topology.

a. Point-to-Point

Point-to-point logical topologies are still common in today's customer premises installations (Fig. 13.15). Two nodes requiring communication are directly linked by the fibers—normally a fiber pair (one to transmit and one to receive).

b. Star

An extension of the point-to-point topology is the logical star topology (Fig. 13.16). This is a collection of point-to-point topology links, all of which have a common node that is in control of the communications system. Common star applications include

- A switch such as a PBX, asynchronous transfer mode, or data switch
- A security video system with a central monitoring station
- An interactive video conference system serving more than two locations

c. Ring

In this topology, each node is connected to its adjacent nodes in a ring (Fig. 13.17). The logical ring topology, especially prevalent in the data communications area, is supported by two primary standards:

- Provides Token Ring (IEEE 802.5)
- FDDI (ANSI X3T9)

d. Logical Bus

The logical bus topology is also utilized by data communications and is supported by the IEEE 802.3 standards. All nodes share a common line rather than in one direction, as on a ring. When one node transmits, all the other nodes receive the transmission at approximately the same time. The most popular systems requiring a bus topology include Ethernet and Manufacturing Automation Protocol.

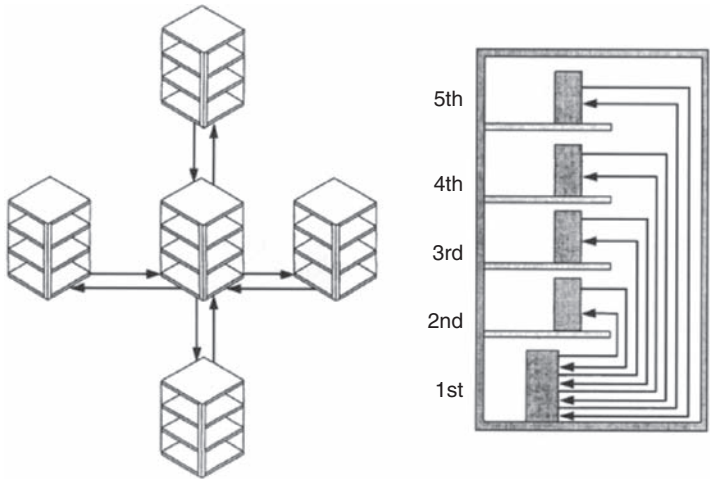


Figure 13.16 Logical star topology.

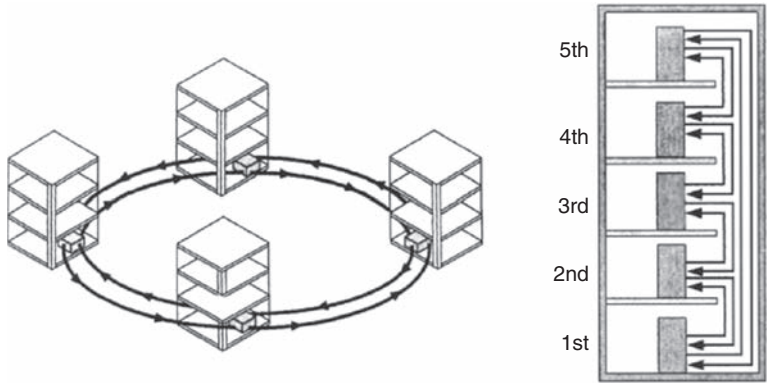


Figure 13.17 Logical counter-rotating ring topology.

e. Token Bus

A token bus topology is shown in Fig. 13.18.

13.3.2.4. Physical Topologies and Logical Topology Implementation

All the logical topologies noted earlier are easily implemented with a physical star cabling scheme, as recommended by TIA/EIA-568-A, Commercial Building Telecommunications Cabling Standard. Implementing point-to-point and star topologies on a physical star is straightforward.

The use of data networks that utilize bus or ring topologies is very prominent in the market, (i.e., Ethernet, Token Ring, and FDDI). With the benefits of physical star

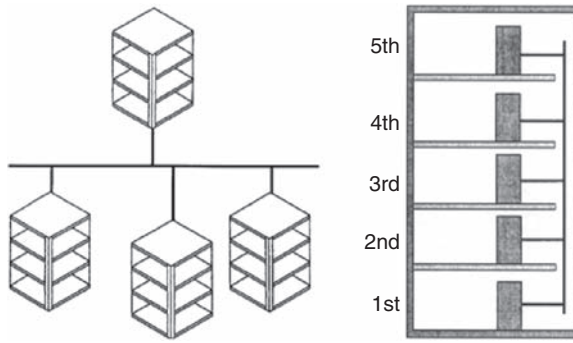


Figure 13.18 Logical bus topology.

cabling, electronic vendors and standards have developed electronic solutions designed to interface with a star network. These applications are typically implemented with an "intelligent hub" or concentrator. This device, in simple terms, establishes the bus or ring in the back plane of the device, and the connections are made from a central location(s). Therefore, from a physical connection, these networks appear to be a star topology and are quite naturally best supported by a physical star cabling system.

13.3.2.5. Physical Stars vs. Physical Rings

In some situations, a physical ring topology would seem appropriate. As seen in the following comparison of an optical fiber physical star topology to a physical ring topology, a physical star topology is recommended for supporting the varied requirements of a structured cabling system.

Physical star

Advantages

- Flexible—supports all applications and topologies
- Acceptable connector loss with topology compatibility
- Centralized fiber cross-connects, making administration and rearrangement easy
- Existing outdoor ducts often configured for the physical star, resulting in easy implementation
- Easily facilitates the insertion of new buildings or stations into the network
- Supported by TIA/EIA-568-A, Commercial Building Telecommunications Cabling Standard

Disadvantages

- Cut cable, resulting in node failure unless redundant routing is used
- More fiber length used in a logical ring topology

Physical ring

Disadvantages

- Less flexible
- Unacceptable connector loss if star or bus logical topologies are required
- Connectors located throughout the system, making administration and rearrangement difficult
- Physical implementation difficult, often requiring new construction
- Insertion of new buildings or stations into a physical ring causing disruptions to current applications or inefficient cable use
- Even with ring topology applications, the majority of nodes, if not all, required to be present and active
- Node survivability in the event of cable cut, if cable truly utilizes redundant paths
- Less fiber lengths used between nodes

13.3.2.6. Physical Star Implementation

The more important recommendations of the TIA/EIA-568-A standard as they relate to optical fiber are summarized as follows: The standard is based on a hierarchical star for the backbone and a single star for horizontal distribution (Fig. 13.19).

The rules for backbone cabling include a 2000-m (3000 m for single mode) maximum distance between the MC and the HC for 623,125 pm and a maximum of one IC between any HC and MC. The MC is allowed to provide connectivity to any number of HCs.

The standard does not distinguish between interbuilding (outside) or intrabuilding (inside) backbones because these are determined by the facility size and

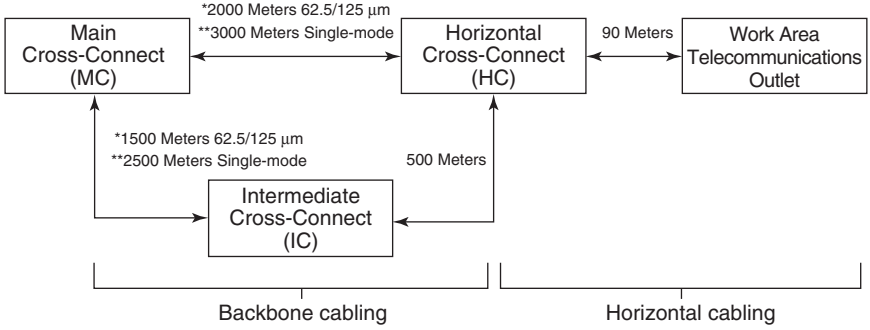


Figure 13.19 Commercial Building Telecommunications Cabling Standard (TIA-EIA-568-A). Backbone distances based on optical fiber cable. When IC to HC distance is less than 500 meters, then the MC to HC shall not exceed 2000 m for 62.5/125 μm or 3000 m for single-mode. Single-mode fiber can support distances up to 60 km (37 miles); however, this is outside the scope of TIA/EIA-568A.

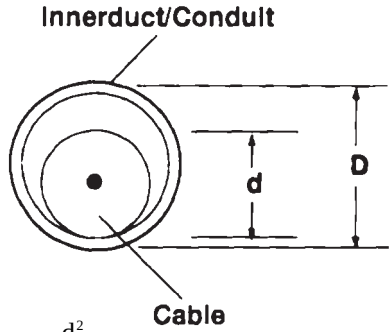


Figure 13.20 Duct utilization: $\frac{d^2}{D^2} < 50\%$, where *d* is the cable diameter and *D* is the innerduct diameter.

campus layout. However, in most applications, one can envision the MC to IC being an interbuilding backbone link. The exceptions to this would be major high-rise buildings in which the backbone may be entirely inside the building or a campus containing small buildings with only one HC per building, therefore eliminating the requirement for an IC (Fig. 13.20).

The horizontal cabling is specified to be a single-star linking horizontal cross-connect closet to the work-area telecommunications outlet with a distance limitation of 90 m. This distance is based not on fiber capabilities but on copper distance limitations to support data requirements. Currently, TIA is studying two alternate horizontal cabling topologies when using fiber in the horizontal: the multiuser outlet and centralized optical fiber cabling.

13.3.3. TIA/EIA-598-A, Optical Fiber Cable Color Coding

A common denominator for all cable designs is identification/color coding of fibers and fiber units. TIA/EIA-598-A defines identification schemes for fibers, buffered fibers, fiber units, and groups of fiber units within outside-plant and premise optical fiber cables. This standard has been adopted by the Rural Utility Service within 7 CFR 1755.900 and Insulated Cable Engineers Association, Incorporated S-87–640-1992, Standard for Fiber Optic Outside Plant Communications Cable.

This standard allows for fiber units to be identified by means of a printed legend. This method can be used for identification of fiber ribbons and fiber subunits. The legend will contain a corresponding printed numerical position number and/or color for use in identification (Table 13.1).

Table 13.1
Color Code.

Position Number	Base Color and Tracer	Abbreviation
1	Blue	BL
2	Orange	OR
3	Green	GR
4	Brown	BR
5	Slate	SL
6	White	WH
7	Red	RD
8	Black	BK
9	Yellow	YL
10	Violet	VI
11	Rose	RS
12	Aqua	AQ
13–24	Colors 1–12 repeated with black tracer	BL/BK, OR/BK etc

13.4. HANDLING AND INSTALLING FIBER OPTICS

13.4.1. Cable Handling Considerations

13.4.1.1. Minimum Bend Radius

Optical fiber cable installation is as simple as, and in many cases much easier than, installing coaxial or *UTP* cable in the horizontal. The most important factor in optical fiber cable installations is maintaining the cable's minimum bend radius. Bending the cable tighter than the minimum bend radius may result in increased attenuation and broken fibers. If the elements of the cable are not damaged, when the bend is relaxed, the attenuation should return to normal. Cable manufacturers specify minimum bend radii for cables under tension and for long-term installation (Table 13.2).

13.4.1.2. Maximum Tensile Rating

The cable's maximum tensile rating must not be exceeded during installation. This value is specified by the cable manufacturer. Tension on the cable should be monitored when a mechanical pulling device is used. Hand pulls do not require monitoring. Circuitous pulls can be accomplished through the use of backfeeding or centerpull techniques. For indoor installations, pull boxes can be used to allow cable access for backfeeding at every third 90° bend (Table 13.3).

Table 13.2
Minimum Bend Radii.

		Minimum Bend Radius			
		Loaded		Unloaded	
Application	Fiber Count	cm	in.	cm	in.
Interbuilding backbone	2–84	22.5	8.9	15.0	5.9
	86–216	25.0	9.9	20.0	7.9
Intrabuilding backbone	2–12	10.5	4.1	7.0	2.8
	14–24	15.9	6.3	10.6	4.2
	26–48	26.7	10.5	17.8	7.0
	48–72	30.4	12.0	20.3	8.0
	74–216	29.4	11.6	19.6	7.7
Horizontal cabling	2	6.6	2.6	4.4	1.7
	4	7.2	2.8	4.8	1.9

Table 13.3
Maximum Tensile Load.

		Maximum Tensile Load			
		Short Term		Long Term	
Application	Fiber Count	N	Ibs	N	Ibs
Interbuilding backbone	2–84	2700	608	600	135
	86–216	2700	608	600	135
Intrabuilding backbone	2–12	1800	404	600	135
	14–24	2700	608	1000	225
	26–48	5000	1124	2500	562
	48–72	5500	1236	3000	674
	74–216	2700	600	600	135
Horizontal cabling	2	750	169	200	45
	4	1100	247	440	99

Note: Specifications are based on representative cables for applications. Consult manufacturer for specific information.

13.4.1.3. Maximum Vertical Rise

All optical fiber cables have a maximum vertical rise that is a function of the cable's weight and tensile strength. This represents the maximum vertical distance the cable can be installed without intermediate support points. Some guidelines for vertical installations include the following:

- All vertical cable must be secured at the top of the run. A split mesh grip is recommended to secure the cable.
- The attachment point should be carefully chosen to comply with the cable's minimum bend radius while holding the cable securely.
- Long vertical cables should be secured when the maximum rise has been reached.

13.4.1.4. Cable Protection

If future cable pulls in the same duct or conduit are a possibility, the use of innerduct to sectionalize the available duct space is recommended. Without this sectionalization, additional cable pulls can entangle an operating cable and could cause an interruption in service. Care should be exercised to ensure the innerduct is installed as straight as possible, without twists that could increase the cable pulling tension.

When the cable is installed in raceways, cable trays, or secured to other cables, consideration should be given to movement of the existing cables. Although optical fiber cable can be moved while in service without affecting fiber performance, it may warrant protection with conduit in places exposed to physical damage.

13.4.1.5. Duct Utilization

When pulling long lengths of cable through duct or conduit, less than a 50% fill ratio by cross-sectional area is recommended. For example, one cable equates to a 0.71411 outside-diameter cable in a l-in. inside-diameter duct. Multiple cables can be pulled at once as long as the tensile load is applied equally to all cables. Fill ratios may dictate higher fiber counts in anticipation of future needs. One sheath can be more densely packed with fiber than multiple cable sheaths. In short, for customer premises applications, the cost of extra fibers is usually small when these extra fibers are not terminated until needed. For a difficult cable pull, extra fibers installed now but not terminated may be the most cost-effective provision for the future (Fig. 13.20).

13.4.1.6. Preconnectorized Assemblies

Of special consideration are preconnectorized cables. Although the use of factory-terminated cross-connect and interconnect jumper assemblies is acceptable, the use of preconnectorized backbone and distribution cable presents special installation techniques. These connectors must be protected when installing the connectorized end of these cables. Protective pulling grips are available to protect connectors, but the grips' outside diameter may prevent installation in small innerducts or conduits. The size of the preconnectorized assembly and pulling grip should be considered before ordering factory connectorized cables. Additional

installation requirements may also be imposed on the grip by the manufacturer, in terms of minimum bend radius and tension, that would be the limiting parameters in an installation.

13.4.1.7. Slack

A small amount of slack cable (20–30 ft) can be useful in the event that cable repair or relocation is needed. If a cable is cut, the slack can be shifted to the damaged point, necessitating only one splice point in the permanent repair rather than two splices if another length of cable is added. This results in reduced labor and hardware costs and link loss budget savings.

Additional cable slack (approximately 30 ft) stored at planned future cable drop points will result in savings in labor and materials when the drop is finally needed. Relocation of terminals or cable plant can also take place without splicing if sufficient cable slack is available.

13.4.2. Cable Splicing Methods

In the commercial building and campus environment, the designer/installer can often avoid the requirement of fiber-to-fiber splicing by installing a continuous length of cable. This is normally the most economical and convenient solution. However, because of the cable plant layout, length, raceway congestion, or requirements to transition between nonlisted and UL-listed cable types at the building entrance point, splices cannot always be avoided. This requires splicing cables together, of which there are two basic techniques. Field splicing methods for optical fibers can be grouped into two major categories: fusion and mechanical.

Fusion splicing consists of aligning the cores of two clean (stripped of coating), cleaved fibers and fusing the ends together with an electric arc. The fiber ends are positioned and aligned using various methods. Alignment can be performed using fixed V-groove, profile alignment, or light injection. It can be manual or automatic and is normally accomplished with the aid of a viewing scope, video camera, or a specialized type of optical power meter for local injection and detection of lights. High-voltage electrodes, contained in the splicer, arc across the fiber ends as the fibers are moved together, thus fusing the fibers together.

A mechanical splice, by comparison, is an optical junction, where two or more optical fibers are aligned and held in place by a self-contained assembly approximately 2 in. in length. Single-fiber mechanical splices rely on the alignment of the outer diameter of the fibers, making the accuracy of core/cladding concentricity critical to achieving low splice losses. This method aligns the two fiber ends to a common centerline, thereby aligning the cores. The cleaned (stripped of coating) fiber ends are cleaved, inserted into an alignment tube, and butted together. The tube has factory installed index-matching gel to reduce reflections and loss at the splice point. Usually, the fibers are held together by compression

or friction, although some older methods rely on epoxy to permanently secure the fibers.

13.4.3. Hardware

The selection of proper hardware for a commercial building or campus environment is essential to provide a flexible cabling system that will conform to the TIA/EIA-598-A, Commercial Building Telecommunications Cabling Standard. The hardware selection process can be narrowed down to a few products based on the following four fundamental categories surrounding the connecting hardware application:

- Indoor or outdoor environment
- Field connectorize or splicing pigtails
- Wall- or rack-mounted
- Fiber count

13.4.4. Indoor Hardware

Indoor hardware selection is complex because of the variety of applications and topologies encountered. Selecting the proper indoor hardware will ensure a flexible fiber-optic cable plant. Indoor hardware can be divided into the application area's MC. The MC is the hub of the cable plant. Consequently, it is the main administrative point and is characterized by a high concentration of fibers as well as multiple indoor and outdoor cables. Adds, moves, and charges are also frequent. Hardware must have a high termination density and capacity. The capacity should be modular, however, to provide solutions that can grow as additional capacity is required. Different termination methods may be employed. Typically, indoor cable is terminated by field connectorization, outdoor cable is terminated by pigtail splicing, and interconnection cables (jumpers) are terminated by factory termination. The hardware must be readily compatible with the different methods, but must be flexible to provide only necessary components for the method employed. The MC generally requires rack-mounted hardware and equipment. Hardware should be wall-mounted if rack space is limited or not available. As the main administrative point, this area will contain a high number of jumpers for both cross-connection of backbone fibers and interconnection to the electronics. Hardware must provide ample room for jumper storage as well as neat, secure routing throughout the rack.

13.4.4.1. Intermediate Cross-Connect

This is the second hierarchical level of cross-connects in the backbone wiring. The IC is generally where the interbuilding backbone and intrabuilding backbone meet. In general, the IC has lower fiber counts than the MC. The required number

of jumpers for cross-connection of backbone fibers and interconnection to the electronics is reduced. Racks with fiber routing provisions may be unnecessary.

13.4.4.2. Telecommunications Closet

This is the location where the backbone wiring and horizontal wiring meet. Typically, this is the location in the network where the fiber backbone transitions to the horizontal distribution to the work areas. The backbone wiring is typically low-fiber count indoor cable. Thus, low-capacity termination housings are usually required. The standard termination method for indoor cables is field connectorization. Thus, only path panels are necessary; the purchase of a splice housing is rarely, if ever, required. The selection of wall- or rack-mountable hardware depends almost entirely on the architectural spaces provided. Jumpers will be required only to interconnect at the electronics. Jumper routing and storage capacities are not as critical.

13.4.4.3. Work Area Telecommunications Outlet

This is the end point of the horizontal wiring. Again, the cables may be fiber only or a combination of fiber and copper. Although the work area usually requires only one outlet, the outlet must usually accommodate two or three different types of media. Mounting is usually in fixed-wall offices or in open-systems furniture; however, floor outlets are sometimes used. Hardware must be compatible with the most common mounting applications and allow a variety of cable entry locations. The work area is a high-traffic zone and therefore hazardous to the cable, connectors, and jumpers. Units must be low profile to minimize the risk of damage. The hardware must provide adequate routing and storage space so that the minimum bend radius of the fiber is maintained.

Outdoor hardware consists of a line of splice closures, wall-mountable distribution centers, and pedestal-mountable cross-connects. These units provide environmental protection for splices, connectors, and jumpers in the outside plant environment, often required in industrial and other special applications.

In some indoor circumstances, space is limited for mounting hardware. Specially designed furcation (or fan-out) kits provide protection and pullout strength for bare fibers, and they are direct connectorized. These are most useful when the fiber counts are low and all of the fibers will be patched into other hardware or electronics in the same area.

13.4.5. Connectors

13.4.5.1. Field and Factory Connectorization

With the advent of easy-to-install field connectors and new methods to fan-out loose-tube cables, field connectorization has become the common method for

terminating fiber-optic cables in the premises market. There are numerous types of connectors on the market, each with slightly different installation procedures.

With regard to all connectors, the fiber must be epoxied into the connector. The epoxy keeps fiber movement over temperature at a minimum, allows polishing without fracturing the fiber, and seals the fiber from the effects of the environment. In addition, it allows the fiber to be aggressively cleaned on the end face.

In addition, the connector end face must be polished. A physical contact (PC) finish is recommended and is specified by TIA/EIA-598-A. This means that the fibers will be physically touching inside the connector adapter as they are held under compression. Lack of a PC finish results in an air gap between the fibers and increased attenuation.

Several polishing methods are recommended, which are typically dependent on the ferrule material used. If a ferrule material is very hard, such as ceramic, it is common for the ferrule to be pre-radiused. Softer ferrule materials, such as composite thermoplastic or glass in ceramic, may be polished flat. These materials wear away at approximately the same rate as the fiber and can be polished aggressively and still maintain a PC finish.

Heat-cured connectors have the advantage of being a cost-effective way to make cable assemblies or to install in a location where a large number of fibers are terminated at one time. Heat-cured connectors typically require more time for the epoxy to harden and skill to polish, and they need epoxies that have limited pot lives. Therefore, they are best used in a controlled environment or where large numbers of connectors are done to take advantage of the low connector costs and special ovens. Connectorized assemblies are almost exclusively heat-cured connectors (Fig. 13.21).

UV-cured connectors can be termed glass-insert connectors. The ferrule features a glass insert surrounded by ceramic. The glass insert allows the installer to do two things. First, the glass insert propagates light. Therefore, a UV-curable adhesive can be used to bond the fiber into the ferrule in a mere 45 s. Second, the glass insert protrudes beyond the ceramic outer sleeve. This means the glass insert is polished along with the fiber. The glass is approximately the same hardness

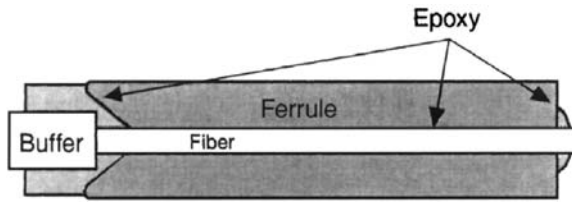


Figure 13.21 Heat-cured connector.

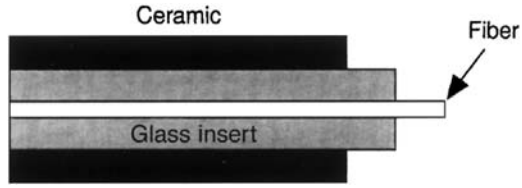


Figure 13.22 Glass-insert connector.

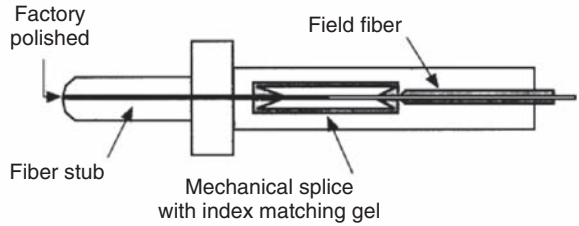


Figure 13.23 No-cure, no-polish connector.

as the fiber and polishes at the same rate. This results in a flat PC polish (Fig. 13.22).

No-cure, no-polish connectors have the advantage of no epoxy, no polish in the field, no consumables, few tools needed, and minimal setup required. These types of connectors are actually a mini-pigtail housed in a connector body. There is a fiber stub already bonded into the ferrule in the factory, where the end face of the ferrule is polished to a PC finish. The other end of the fiber is cleaved and resides inside the connector. The field fiber is cleaved and inserted into the connector until it "butts up" against the fiber stub. A simple cam actuation process completes the connector with no epoxy or polishing required (Fig. 13.23).

REFERENCES

1. Englebert, J., S. Hassett, and T. Summers. *Optical Fiber and Cable for Telecommunications.* In *The Electroncis Handbook*, ed. J. Whitaker, Chapter 10. Boca Raton, Fla.: CRC Press.
2. IBM. 1996. *FTS direct attach physical and configuration planning*, 2nd ed. New York: IBM Corp.
3. Siecor Corporation. 1995. *Universal transport system design guide, release III.* Hickory, N.C.: Siecor.
4. Siecor Corporation. 1996. *Premises fiber optic products catalog*, 6th ed. Hickory, N.C.: Siecor.

This page intentionally left blank

14

Test and Measurement of Fiber Optic Transceivers

Greg D. Le Cheminant
Agilent Technologies, Digital Signal Analysis Division
Santa Rosa, California

The basic fiber-based communications system consists of a laser transmitter which converts electrical data to intensity modulated light, a fiber optic cable as short as a few meters or as long as 100+ kilometers, and a photodetector receiver that converts the information on the light back to an electrical signal.

An indicator of how well the entire system performs is a measurement called bit-error-ratio (BER) which indicates how many bits were received in error compared to the total number of bits transmitted. Acceptable BER's are in the range of one error per one billion to one trillion bits transmitted. It is rare for a newly designed system to operate at these performance levels at initial turn on. Prior to system construction the individual elements must be characterized to determine their potential for proper system level performance. This chapter will focus on testing of the transmitter and the receiver.

When a transmitter is paired with a receiver (through a fiber), and the desired BER is not achieved, is the transmitter at fault, or is it the receiver? Perhaps both are faulty. A poor quality receiver can be compensated for by a high-performing transmitter (and vice versa). When a standard (Ethernet, Fibre Channel etc.) communications system is designed, device specifications are determined so that the cost/performance burden is properly balanced at both ends of the fiber. Specifications are set so that any receiver will interoperate with the worst case allowable transmitter, and any transmitter will provide a signal with sufficient quality such that it will interoperate with the worst case allowable receiver.

14.1. TRANSMITTER TEST

To determine how well a transmitter carries information, the time-domain waveform is analyzed. Most optical communication waveforms are viewed using

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

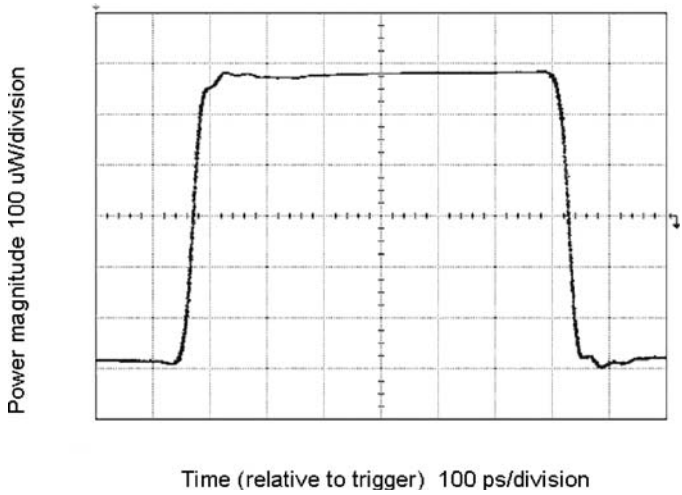


Figure 14.1 Data waveform in the pattern display.

a wide-bandwidth sampling oscilloscope. Two schemes are used to observe the waveform. In general, how the oscilloscope is triggered results in waveforms being displayed as either an eye-diagram or as a data pattern. In the data pattern display, the bits of the signal are shown in a sequential format. Figure 14.1 shows a typical digital communications signal displayed as a data pattern while Figure 14.2 shows the same signal displayed as an eye-diagram.

14.1.1. Construction of the Data Pattern

To observe a data pattern waveform, a special trigger signal, often called a pattern trigger is required. A pattern trigger is typically a signal edge that is generated to coincide with a specific location or position in the data pattern being observed. The pattern trigger is generated once for every repetition of the data pattern. The first sample is taken once the oscilloscope is armed and the trigger edge (often indicating the start of the pattern) triggers the oscilloscope. The pattern trigger must always occur at the same exact time location relative to the data sequence. The oscilloscope will rearm and wait for the next pattern trigger edge. The trigger occurs at the same point in the data sequence. The second sample is taken at a point in the pattern adjacent to and slightly later in time than the first point. The process is repeated over and over again, with the incremental delay increasing and forcing the sampling to take place across the data pattern.

It is important to note the two critical requirements to observe the pattern waveform:

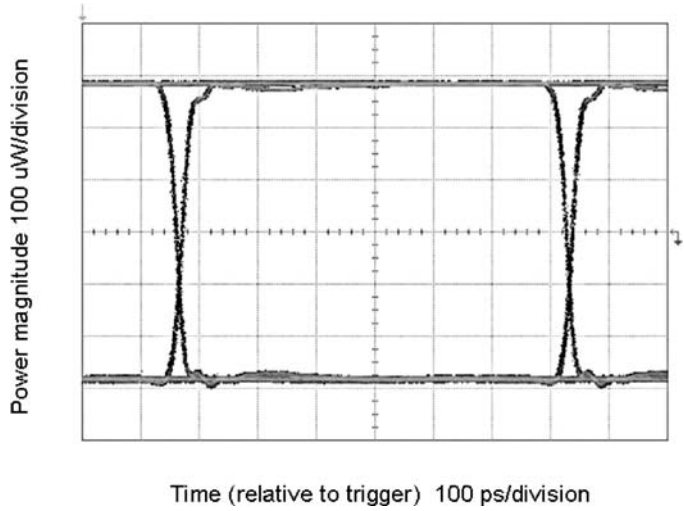


Figure 14.2 Data waveform in the eye diagram display.

1. The data pattern or waveform must be a *repeating signal*
2. The oscilloscope must be triggered by a signal that occurs only once for N repetitions of the pattern (N = 1, 2, 3. . .) and is always at the same relative location to the pattern

Very long patterns present some difficulty for displaying pattern waveforms. Because the trigger pulse is only generated once per every repetition of the pattern, the time between trigger events can become very large. For wide bandwidth sampling oscilloscopes, one data point is sampled for every trigger event. If a waveform is composed of 1000 sample points, the entire pattern must be transmitted 1000 times to complete acquisition of the waveform. For long pattern lengths it can take several minutes and even hours to produce a single waveform. Thus pulse trains are usually only displayed when examining relatively short patterns.

14.2. CONSTRUCTION OF THE EYE-DIAGRAM

Another difficulty when examining pulse trains is that only a few bits can be examined at a given time. More bits can be displayed by decreasing the resolution of the oscilloscope time base (more time per division on the time axis), but important details are usually lost due to this reduced resolution. Often when examining a high-speed digital communication signal it is desirable to determine the overall performance of the system for all sequences within the complete data pattern. It would be ideal to see this in one simple display. This can be achieved

through the eye-diagram. The eye-diagram is a composite display of waveform samples acquired throughout the entire data pattern displayed on a common time base. Consider the eight waveforms that can be generated from a three bit sequence (000, 001, . . . 110, 111). If these eight waveforms are all placed on a common amplitude vs. time grid the eye-diagram is displayed.

The eye-diagram sampling process is similar to that of the pattern waveform but there are some important exceptions.

Clock triggers occur much more frequently than pattern triggers.

Generally the oscilloscope will be triggered more frequently with a clock (assuming the pattern repetition rate is slower than the rearm time of the oscilloscope).

The oscilloscope rearm time is asynchronous to the data signal and trigger.

The time at which the oscilloscope will accept triggers will occur at arbitrary locations in the data pattern. Samples are taken with high precision and synchronism to the data. The clock trigger ensures that the sampling is precisely in step with the bits in the pattern. However, *which* specific bits are sampled in the data stream is uncontrolled because the rearm time of the oscilloscope is completely independent of the signal being observed. Thus two adjacent samples will be taken at completely unrelated locations in the data stream. If sample 1 is taken on a logic 1 level, sample 2 is equally likely to take place on a 0 level or a 1 level. The result is that the complete waveform record will have been acquired from many locations in the bit stream (many 1's, many 0's, as well as the transition regions). The synchronous but arbitrary waveform samples are then all displayed on a common timebase. This is the eye diagram.

With the common sampling oscilloscope eye-diagram it is difficult to view the waveform from any individual bit. However, a wealth of information is available regarding the overall performance of the transmitted waveform. If the eye-diagram begins to close horizontally (along the time axis) this is an indication of excessive waveform timing jitter. Slow rise and fall times cause vertical eye closure. Eye closure due to any mechanism presents a significant system level problem because it makes the decision process more difficult for the receiver at the end of the communication system. This is easily observed with the eye diagram but would likely be difficult to discern in the pattern waveform.

Whether a pattern trigger or a clock trigger is used, it is important to remember that the display of a signal can be no better than the signal used to trigger the oscilloscope. If a pristine signal with extremely low jitter is observed with an oscilloscope triggered by a signal with high jitter, the observed signal will have high jitter too. Everything displayed is relative to the trigger signal. Typically trigger signals are of high quality and yield accurate waveforms. Sometimes triggers based on divided clocks with very high divide ratios are not completely synchronized with the signal being measured. The lack of synchronism results in horizontal eye closure.

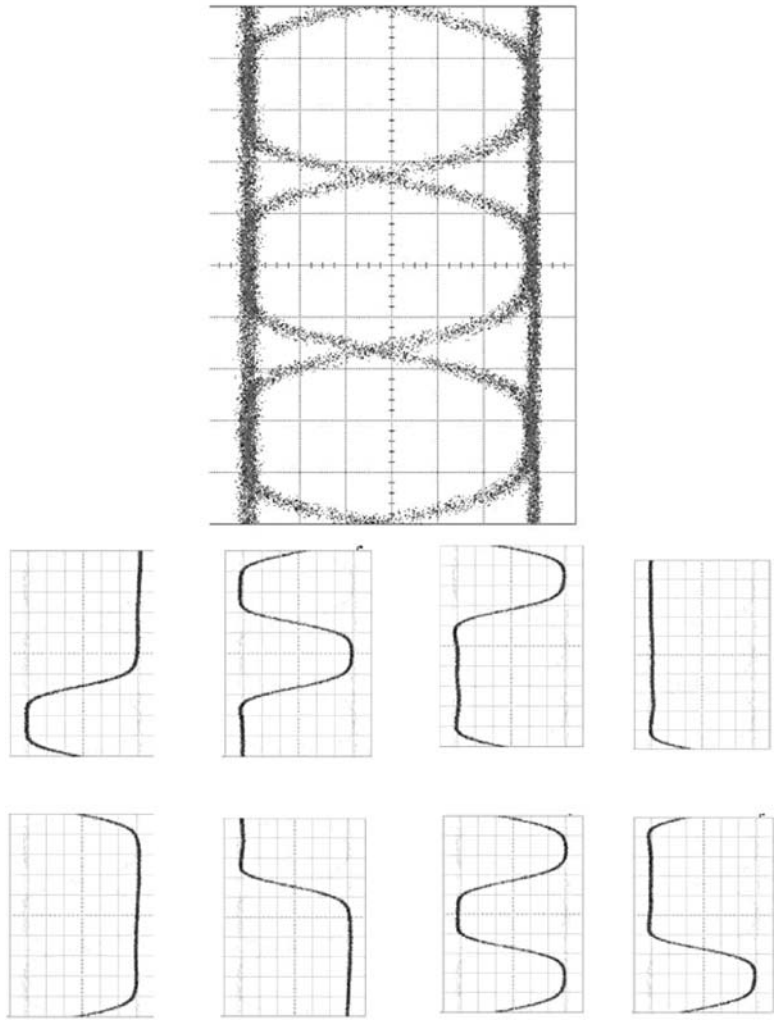


Figure 14.3 Logical construction of the eye-diagram.

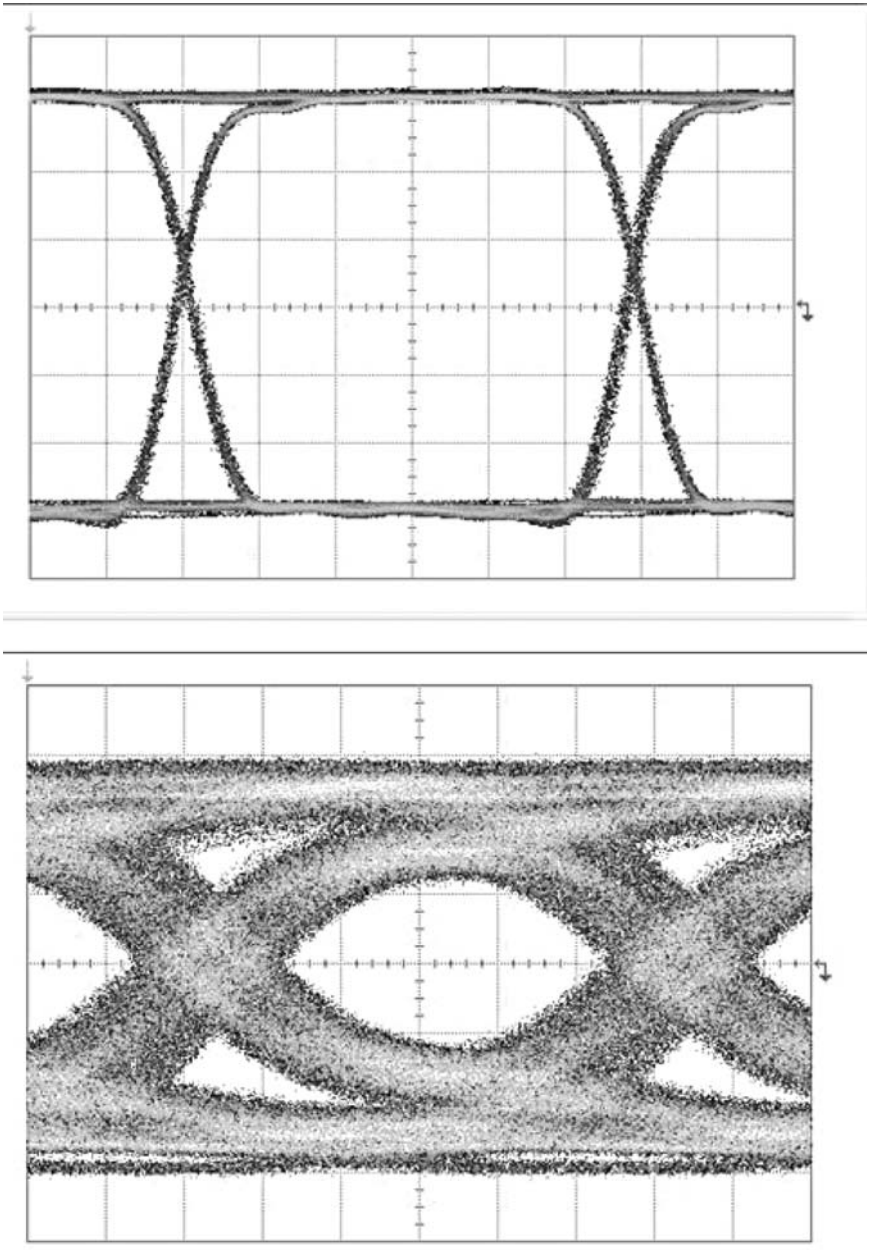


Figure 14.4 Comparing the "ideal" versus the impaired eye diagram.

14.3. OSCILLOSCOPE FREQUENCY RESPONSE

Anytime that an oscilloscope system alters the frequency content of a signal the resulting waveform display will not be an exact replica of the input signal. This can be a significant source of frustration and/or confusion to the user, because there is no indicator or warning from the oscilloscope indicating that the signal is being altered. Perhaps the simplest example of this concept is to examine a very fast edge with a low bandwidth oscilloscope. The high-frequency content of the signal will be suppressed and the observed edge will be slowed. This leads to the general concept that more measurement bandwidth is usually better than less bandwidth. When measuring very fast signals, this is usually a good idea. But there are some tradeoffs and possible problems associated with a very wide-bandwidth oscilloscope, and it is not just that higher bandwidth oscilloscopes are harder to build and cost more money compared to a lower bandwidth instrument.

The internal electronics of the sampler circuits in a wide-bandwidth oscilloscope produce a small amount of noise. This is a thermal noise term that is proportional to the square root of the channel bandwidth. Another possible problem with increasing the oscilloscope bandwidth is much more difficult to recognize and quantify. While a wide bandwidth is essential to avoid suppressing high-frequency elements of a signal, realize that an oscilloscopes' bandwidth specification simply indicates the frequency at which the oscilloscope response drops by 3 dB. That is, if the oscilloscope is used to measure simple sine waves, as the frequency of the sine wave is increased, eventually the displayed magnitude of the sine wave will be less than the actual magnitude. The frequency at which the displayed value is at 70.7% of actual (−3 dB compared to the response at DC) is used to indicate the oscilloscope bandwidth. It does not indicate what the frequency response (the overall attenuation characteristic versus frequency) was before, and just as important, after the −3 dB corner frequency. There are many possible frequency response characteristics that have the same bandwidth as seen in Figure 14.5. If a signal is observed with oscilloscopes having identical bandwidths but different frequency responses, and the signal has significant spectral content in the region of the −3 dB rolloff, the observed signals can have significantly different waveform displays. Remember, if the frequency content of the signal is altered, the time domain signal shape will be changed! Different frequency response characteristics will cause the various frequency elements of a signal to be attenuated in different ways.

An oscilloscope channel response with a gradual rolloff, even with a comparably lower bandwidth, will produce a more faithful output waveform than the oscilloscope with the wider but more abrupt frequency response. Unfortunately, while the bandwidth of the oscilloscope is always specified, the frequency response rarely is. Also, filter design theory also tells us that the phase response of

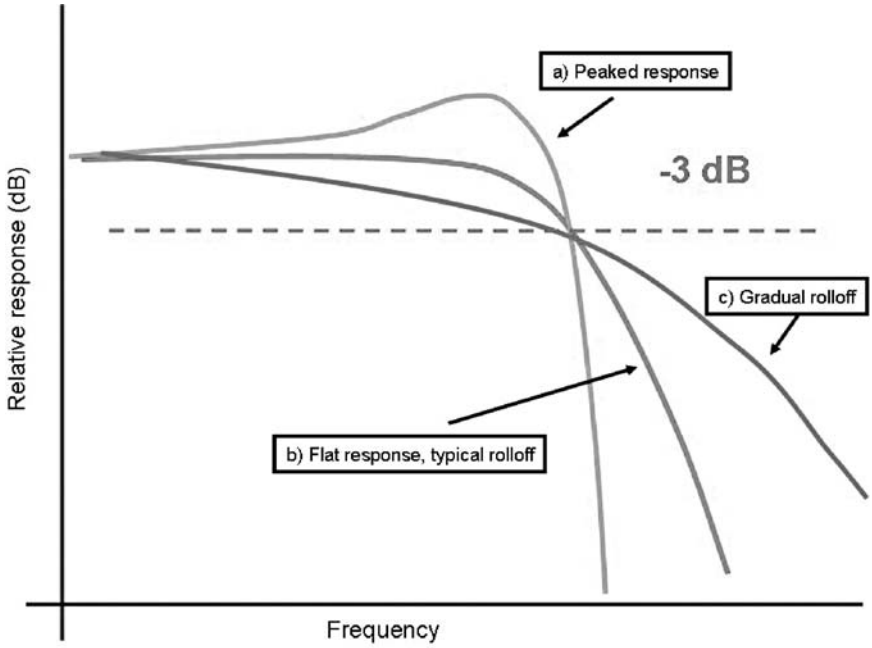


Figure 14.5 The frequency response of the oscilloscope will impact the displayed waveform.

a circuit can have a significant impact on the time domain response. Again, bandwidth, and not phase response of the oscilloscope is likely the only specification that is available.

14.4. EYE MASK TESTING

Good amplitude separation and low jitter are seen as an eye diagram with a wide opening both in time and amplitude. Rather than make several measurements to determine the quality of the eye, it is possible to do this in one test. The openness of an eye diagram can be verified by performing an eye mask test. A mask consists of several polygons that are placed in and around the eye diagram, indicating areas where the waveform should not exist. A "good" waveform will never intersect the mask and will pass the mask test; a "bad" waveform will cross or violate the mask and fail.

14.5. EYE MASK DIMENSIONS AND COORDINATE SYSTEMS

Eye mask shapes are typically defined by industry standards. Virtually every eye mask will have a polygon (keep out region or "no-fly" zone) placed in the

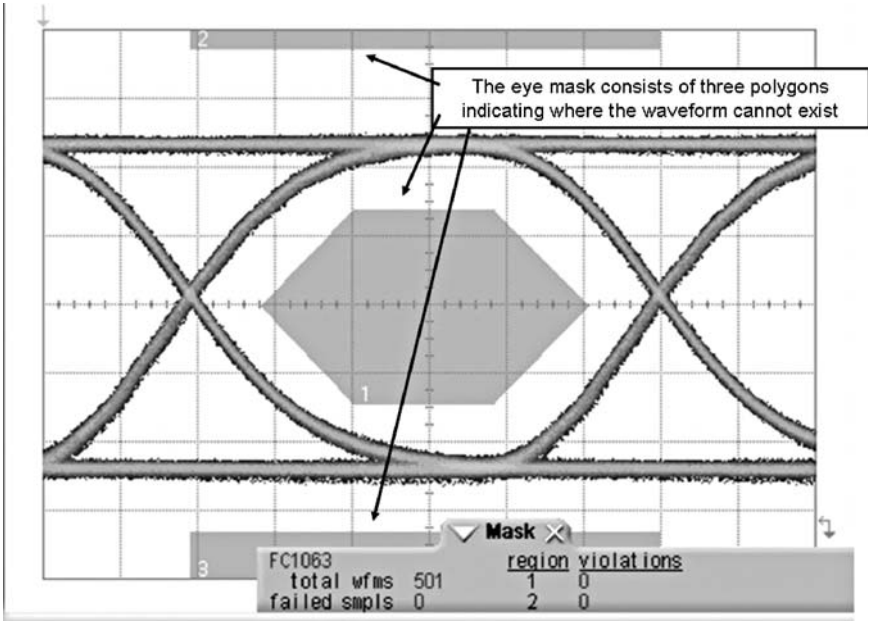


Figure 14.6 Eye mask test for a complaint transmitter.

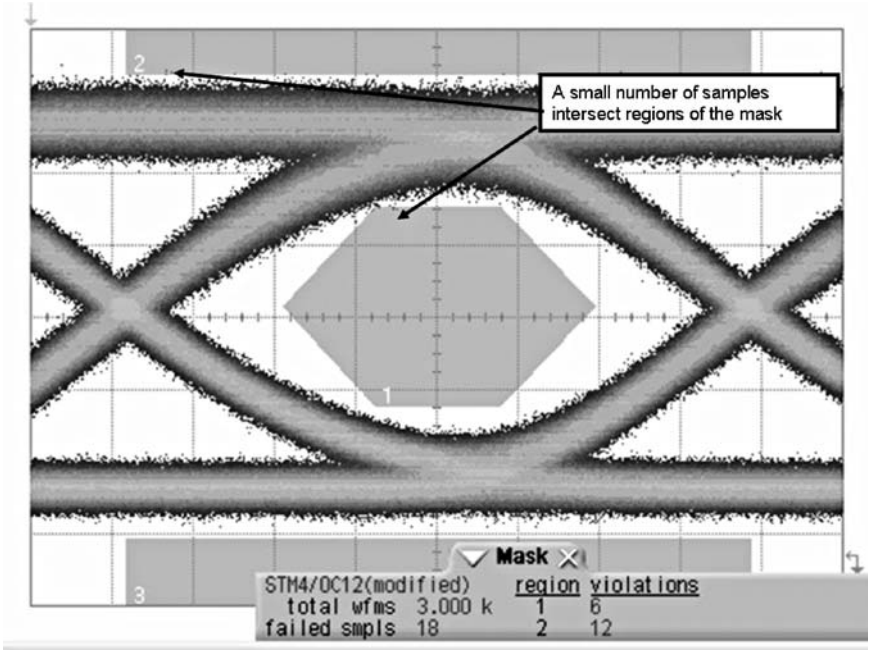


Figure 14.7 Eye mask test for a non-compliant transmitter.

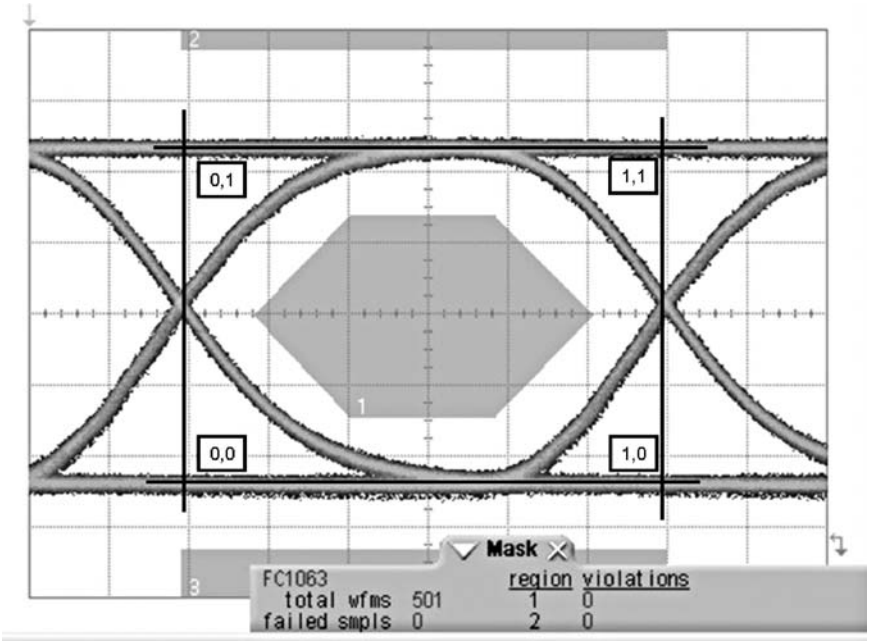


Figure 14.8 Mask coordinate system. Most optical masks are defined in amplitude by the logic 1 and logic 0 levels, and in time by the location of the crossing points of the eye.

center of the eye diagram. Many standards also place a polygon above the logic one level (to screen overshoot in 0 to 1 transitions) and a polygon below the logic 0 level (to screen overshoot in 1 to 0 transitions). Different standards use different shapes for the central polygon. Most mask vertical dimensions are made proportional to the eye amplitude. That is, the mask is scaled vertically according to the signal strength of the eye. The horizontal scaling is generally fixed according to the expected data rate of the signal.

The mask dimensions can be described according to a general rectangular coordinate system. See Figure 14.8. The X axis origin is set at the beginning of the bit period, while one unit later in time represents the end of the bit period. The Y axis origin is defined as the mean logic 0 level. One unit above is set at the mean logic 1 level. Any mask shape can then be defined according to this coordinate system.

14.6. AUTOMATIC MASK TESTING

Performing a mask test takes advantage of many of the analysis tools used for eye diagram parameters. To automatically position the mask polygons in time, a single crossing point is located on the signal presented to the oscilloscope. (Since

the mask dimensions are defined for a specific data rate, knowing the location of one crossing point will allow the oscilloscope to anticipate the location of the next crossing point without making a second crossing point measurement). This defines the x axis of the mask coordinate system. The mean logic 1 and logic 0 levels are measured to define the Y axis of the mask coordinate system. The mask dimensions are then scaled and the mask is placed on the displayed eye diagram.

The oscilloscope will determine if any waveform samples fall on the mask. Most industry standards define a transmitted signal as being non-compliant if *any* waveform samples violate the mask. This brings up an interesting problem with eye mask testing. A larger population of waveform samples should provide a more accurate assessment of the transmitter performance. However, every eye diagram will have random characteristics in both amplitude (noise) and time (jitter). As more data is collected and compared to the mask, the likelihood of mask violations increases. In theory, if enough samples are acquired, eventually almost any transmitter will fail a mask test. The probability for mask violations may be very low, but mask testing generally does not account for probabilities. It is an all or nothing test, pass or fail. This brings up an interesting question. How many samples need to be acquired to perform the test?

At a minimum, there needs to be enough samples to allow the oscilloscope to have sufficient data to align the mask to the waveform. Typically a small population will be sufficient. How much more data should be acquired? If there is significant space (margin) between waveform samples and the mask, it is likely that collecting more data will only slightly change the result. The problem occurs when a device barely passes the mask test. Collecting more samples could lead to a failure.

Some recent industry standards, such as IEEE 802.3 ah (a short reach optical standard) recognize the problem and allow mask failures to occur. However, only a very small ratio of hits to samples is allowed. This allows a large population to be collected, possibly yielding a more accurate measurement without an increased likelihood of failure.

14.7. FINDING THE ROOT CAUSES OF EYE MASK FAILURES

Mask testing can also be automated to take special action when mask violations occur. Waveform acquisition can be halted, the violating waveform can be captured and saved, or tests can be run to capture only a specific number of waveforms regardless of the number of failures. This leads to some interesting modifications of the standard mask test process. Mask failures typically occur when the transmitter waveform is at the extremes of its performance. This can occur when random noise and/or jitter are at extreme levels. Perhaps a more likely

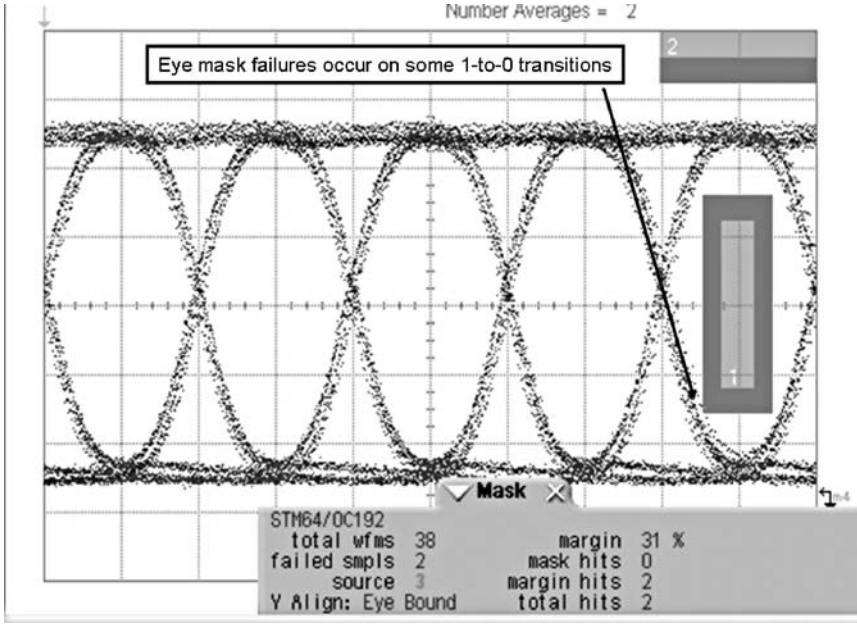


Figure 14.9 Using mask margins to identify stress sequences. The mask size is increased until mask violations occur.

scenario is when some pattern dependent effect is at its extreme. For example, one could imagine that the transient response of a transmitter will be different for a 1010101010 type pattern compared to a 0000000001 pattern. By observing the waveform leading up to a mask violation, these pattern dependencies can be determined.

In the following example (Figs. 14.9 and 14.10), a standard mask test is configured where the waveform will be captured when a mask violation occurs. Normally, no mask violations are observed. However, as the mask margin is increased (increasing the size of the mask beyond the standard dimensions), eventually failures are forced to occur. The "difficult" pattern sequence in this case is a long run of consecutive ones transitioning to a zero.

14.8. EYE MASK SHAPES

Eye masks come in several shapes and sizes depending on the industry standard that defined the specifications for a transmitter. While the upper and lower regions of the mask are generally simple rectangles, the central portion, the mask that is placed inside the eye, can be as simple as a rectangle or as complicated as a 10-sided polygon. One obvious question is how are the shapes and sizes of

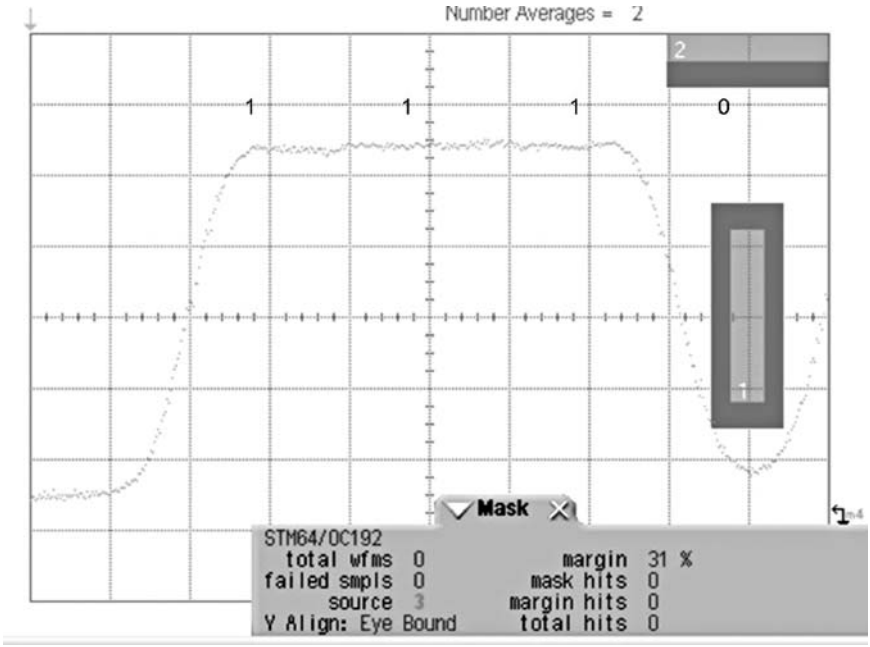


Figure 14.10 The oscilloscope is configured to display the waveform sequence that led to the violation (a 1-1-1-0 pattern).

masks designed? When a new standard is designed, the mask is typically derived from a mask that existed in a previous related standard. In some cases the mask is simply scaled in time to match a new data rate. On the other hand, there have been cases where a mask has been "shaved down" from an existing shape simply because those writing the standard needed some headroom to make sure their devices could pass the test! In general, the central polygon is six-sided as in Figure 7.23. However, a simple rectangle (horizontal extremes removed, SDH/SONET 2.5 and 10 Gb/s)) or clipped down corners to create a 10-sided polygon (Ethernet 10 Gb/s) have also been created.

14.9. THE REFERENCE RECEIVER

As mentioned at the beginning of the chapter, a transmitter needs to be capable of providing a signal with quality sufficient for the worst case allowable receiver. If the receiver's perception of the waveform is critical, then it is important to find a way to do the mask test from the "perspective" of the system receiver. For many years, the concept of a reference receiver has been the foundation of standards based optical transmitter test. A reference receiver is an optical-to-electrical

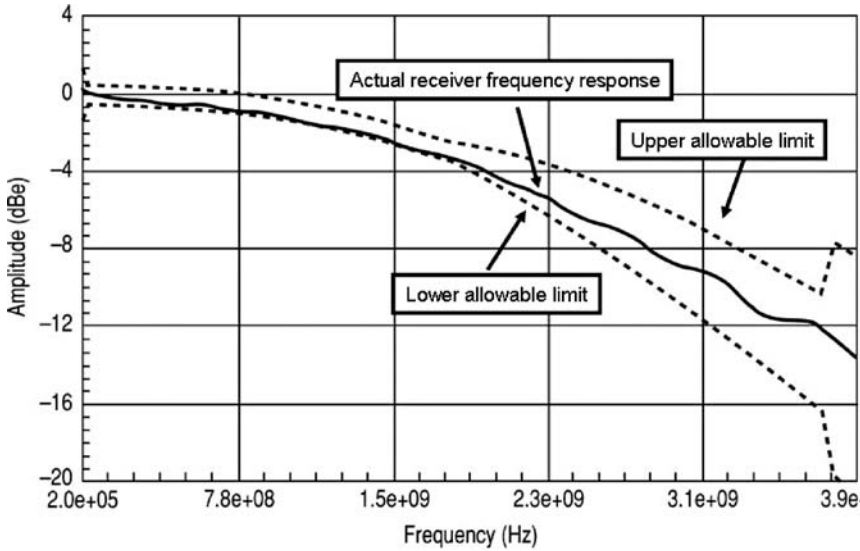


Figure 14.11 Frequency response of a reference receiver for SONET/SDH 2.488 Gb/s.

converter with a fourth order Bessel-Thomson frequency response whose −3 dB bandwidth is set to a frequency of 75% of the transmission data rate. For example, if the transmitter operates at 10 Gb/s, the reference receiver bandwidth will be 7.5 GHz. The reference receiver is used to give the test system the desired communications system receiver perspective. While a system receiver is not required to have a "75% of bit rate" bandwidth, this characteristic is a close approximation of the ideal receiver response for an NRZ waveform.

If the bandwidth of an oscilloscope system attenuates the frequency content of a signal, it will not provide a completely accurate representation of the waveform. In almost all cases, a bandwidth less than the data rate will significantly alter the shape of a digital communications waveform. In Figure 14.12, a laser transmitter operating at 1.25 Gb/s is measured both with a 10 GHz bandwidth optical oscilloscope and the same oscilloscope configured with a 1.25 Gb/s optical reference receiver (938 MHz bandwidth). In this case, a precision filter is switched in the signal path to reduce the 10 GHz bandwidth to 938 MHz). When the signal is viewed in the eye-diagram format with the 10 GHz bandwidth setting, significant overshoot and ringing (a common phenomenon with high-speed laser transmitters) is observed. The overshoot is so severe that the waveform fails the mask test. However, when the oscilloscope is configured as a reference receiver (Fig. 14.13), the high-frequency content of the signal is suppressed, the signal appears to be very well behaved, and the waveform easily passes the mask test.

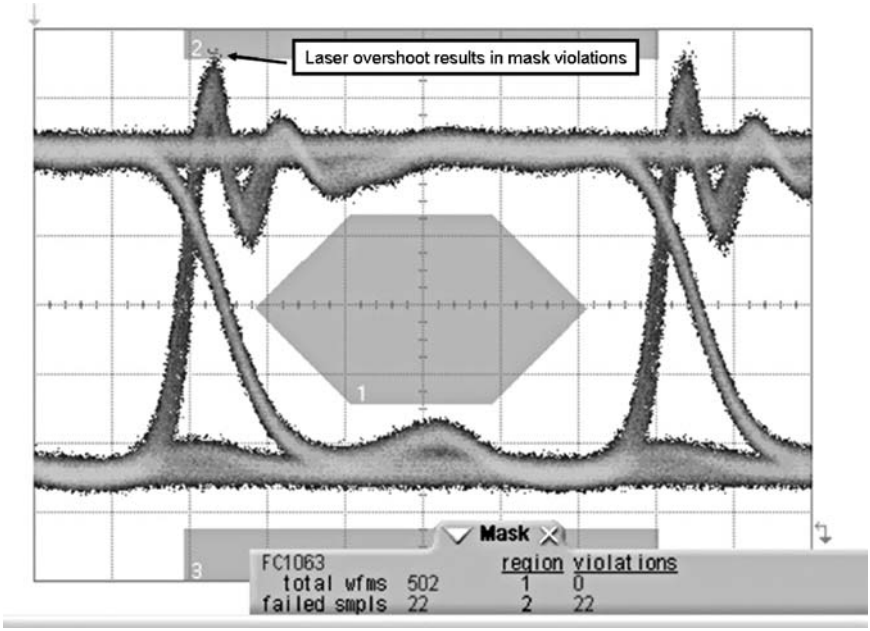


Figure 14.12 Laser eye diagrams with a wide bandwidth oscilloscope having a ~10 GHz bandwidth.

The use of a reference receiver seems counterintuitive, because it appears to have cleaned up the true behavior of the laser and possibly made a bad device look good. This is where it becomes important to take a step back and remember the goal of testing. It is not to precisely characterize the behavior of the laser. Rather, the test is intended to determine how well the laser will interoperate with a receiver in a real communications system. System receivers will not have infinite bandwidth, so it does not make sense to test the transmitter as if it were communicating with a receiver that did. Instead, receivers often have just enough bandwidth to correctly differentiate a logic 1 from a logic 0. For a non-return-to-zero (NRZ) signal the ideal system bandwidth for this will be near 75% of the data rate, the bandwidth used by the reference receiver. Thus the test system reference receiver provides a good representation of the signal from the perspective of a system receiver. Additionally, a reference receiver provides consistency in test. If everyone tests with a specific measurement system bandwidth, results should not vary between test systems. Vendors and their customers should get similar test results, as long as they use agreed upon receiver bandwidths in their test systems.

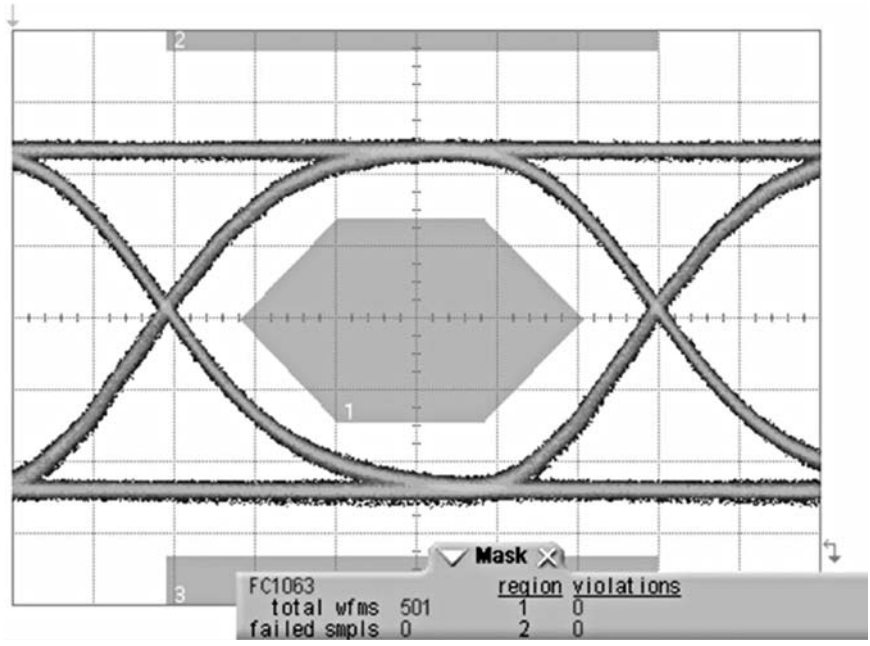


Figure 14.13 The same signal from figure 12 is viewed with a reduced bandwidth reference receiver.

What is the basis for using the specific "75% data rate Bessel-Thomson" design in the test system receiver? First, real *system* receivers will have limited bandwidth to optimize the signal to noise ratio at the decision circuit. It is then desirable to observe the transmitter waveform in a reduced bandwidth also. What should that bandwidth be? The lower the bandwidth, the lower will be the noise. However, if the bandwidth is reduced too much, there will eventually be inter-symbol interference, as the waveform takes longer to transition from one logic level to another. This results in waveform trajectories that begin to close down the eye. The lowest bandwidth that will not result in inter-symbol interference is at 75% of the data rate. This is the "Thomson" element of the filter.

Why use a Bessel type filter design? The Bessel filter design achieves the most well behaved time domain response. That is, while its frequency domain characteristics are inferior to some filter designs in terms of its ability to suppress higher frequency signals, the Bessel design has superior waveform performance with minimal eye diagram distortion (overshoot, ringing etc.). This is because the Bessel filter has constant group delay in the filter passband.

14.10. LASER TRANSMITTER EXTINCTION RATIO

It is important for a transmitter's signal strength to be large enough to overcome the noise present in a system. It is also important for the signal to be strong enough to maintain distinct logic levels after traveling a long distance. Early optical high-speed communications systems often spanned large distances, and required repeaters to overcome channel attenuation. To minimize the number of repeaters and reduce cost, it became important to maximize the information content available given the power available from laser transmitters.

One measure of a laser's communications efficiency is the *ratio* of the 1 level to the 0 level. This measurement is called extinction ratio (ER). Consider a transmitter that sends logic ones at a level of 1 mW and logic zeros at 0.1 mW. Consider another transmitter that sends ones at 1.5 mW and zeros at 0.6 mW. From the receiver perspective, where amplitude separation is critical, both transmitters would perform equally well, because both have a separation between logic levels of 0.9 mW. However, the second transmitter requires significantly higher power to achieve that separation. The first laser has an ER of 10 (1 mW divided by 0.1 mW) and is far more efficient than the second laser, which has an ER of 2.5 (1.5 mW divided by 0.6 mW). The extinction ratio is used as an indicator of how well available laser power is converted to modulation power. High extinction ratios are usually achieved by forcing the zero levels close to a no power state. When extinction ratio is high, virtually all of the available laser power is being converted to information power.

The procedure for measuring extinction ratio has been standardized under TIA OFSTP-4A and IEC 61280-2-2. These procedures specify two things. First, the measurement hardware and configuration is described. Second, the procedure for acquiring and analyzing the data is described. The test hardware for an extinction ratio measurement includes an optical-to-electrical (O/E) converter, a low-pass electrical filter, and a digitizing sampling oscilloscope. The O/E converter transforms the optical signal to an electrical signal that can then be measured by the electrical oscilloscope. This is essentially the reference receiver described above for laser eye-mask test. The oscilloscope/reference receiver combination provides an easy to use method to view the laser waveform.

The filtering in the test system is used for two reasons. First, it provides a consistent measurement system. That is, two test systems with the same filtered frequency response should yield consistent test results, as the extinction ratio measurement is affected by the shape of the eye-diagram. The shape of the eye-diagram can be highly dependent upon the frequency response of the test system. It is important to note that the frequency response is specified for the combination of the O/E converter and the filter, and not the filter alone. A second reason for filtering is that the integrating function of a true receiver is mimicked by this filtering and thus replicates performance in a real system.

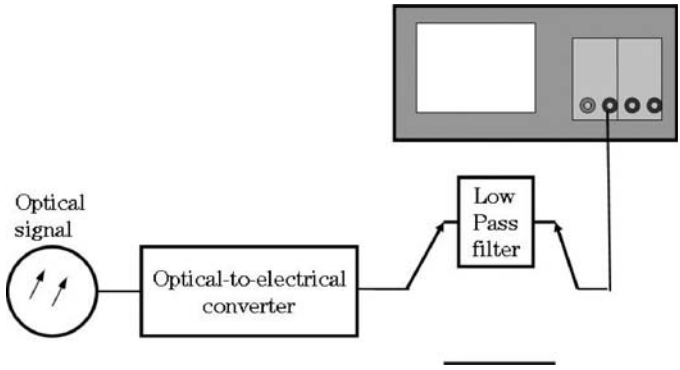


Figure 14.14 Equipment setup for extinction ratio. The optical-to-electrical receiver and low-pass filter are typically integrated into the oscilloscope to minimize signal degradation caused by cabling and connectors that can occur with external configurations. The filter can also be bypassed for full bandwidth analysis.

Although the filter and photodetector are shown as being external to the oscilloscope, for measurement integrity reasons, these components should be integrated into the measurement system and confirmed to meet the specifications for a reference receiver. As data rates move well into the high Gb/s range, designing and building a compliant reference receiver is a complicated "RF/microwave" engineering task. The entire measurement path from the optical connector to the oscilloscope display must meet the tight reference receiver frequency specification. Attaching a well designed electrical filter to a wide bandwidth photodetector (as separate components), and then the pair to the front end of an oscilloscope may come close to achieving a reference receiver response, but in practice the combination is rarely compliant and the test results are often corrupted.

Once the eye diagram is generated, histograms are used to do a statistical analysis of the waveform. The histogram is analyzed to determine the mean "1" level and mean "0" level. From these two values, extinction ratio can be calculated. For example, if the mean level of the upper histogram yields a "1" level of 59.97 and the mean level of the lower histogram yields a "0" level of 3.313, taking the ratio of the two values yields an extinction ratio of 18.1. To compute a decibel value for extinction ratio, simply take the base 10 logarithm of the ratio and multiply by 10 to get 12.6 dB. This level of extinction ratio is common for a laser operating at 2.5 Gbit/s (SONET OC-48/SDH STM-16).

14.11. EXTINCTION RATIO MEASUREMENT ACCURACY

When any measurement is made it is always wise to question how accurate the results are. There are many sources of uncertainty for the extinction ratio

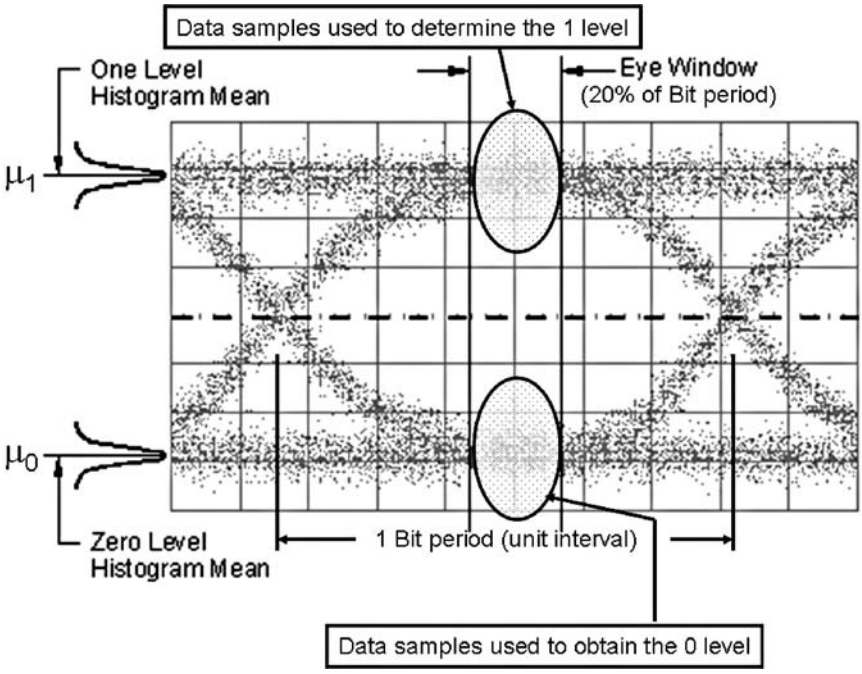


Figure 14.15 Using histograms to compute extinction ratio. The data is acquired in the center of the eye, similar to where a receiver will make its decision on the bit level in an actual communications system.

measurement. It will be apparent that the extinction ratio calculation can be influenced by several factors that are difficult to detect by simply looking at the eye diagram display. These can be grouped into three categories:

- Offsets and spurious signals generated by the instrumentation
- Waveform distortion caused by the instrumentation
- Precision of the measuring instrument

Getting the best accuracy from an extinction ratio measurement requires well designed hardware and a measurement routine that is able to compensate for systematic sources of error.

One of the largest potential sources of measurement error is from "dark signals" or "dark levels". A dark level is a signal that is present even when there is no optical power being fed to the instrument. These can be generated by photodiode dark currents or offset voltages from electrical amplifiers (in the O/E receiver chain). These signals have the effect of offsetting the eye diagram in either a negative or positive direction. Both the "1" level and "0" level will be shifted. Because the offset signal *adds* to both "1" and "0" levels, it does not "common

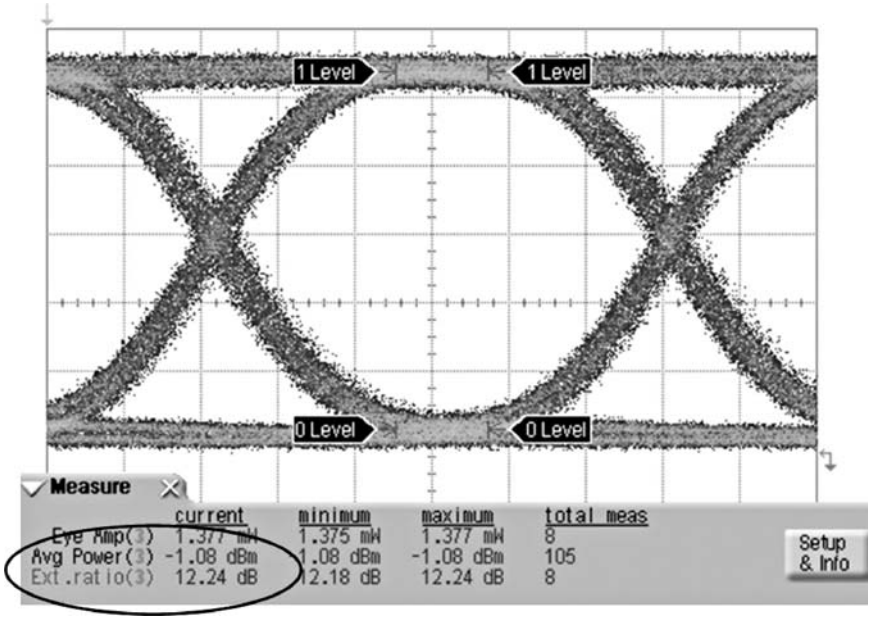


Figure 14.16 A digital communications analyzer ER measurement.

mode out". The ER measurement result will be altered. However, the dark level is a systematic error mechanism. This allows it to be removed from the measurement result.

To determine what the dark levels are in an extinction ratio measurement, simply block the input to the optical receiver and measure what signal is present. (It is important to measure the dark level at the same oscilloscope vertical scale setting that will be used when the actual extinction ratio measurement is made. Dark level can be influenced by the vertical scale setting.)

When the dark level is observed, it typically will be small and may appear to be insignificant. However, keep in mind that small errors in measuring the "0" level can lead to large errors in the ER result. Thus any dark level offset should be accounted for. With the dark level measured, the extinction ratio measurement is adjusted by simply removing the offset it causes from both the "1" and "0" levels. Dark level measurement and effective removal are part of an automatic calibration process built into most modern digital communications analyzers based on equivalent time sampling oscilloscopes. It should be clear that accurate removal of the dark level requires that the dark level be stable. If it is not, the dark level will have to be continually re-measured. Well-designed instrumentation will typically have both small as well as stable dark levels.

14.12. OPTICAL MODULATION AMPLITUDE (OMA)

As laser transmitter technology improved, lower cost transmitters became available, and high-speed optical local area networks (LAN) have become practical. Due to the relatively short spans of such systems, laser efficiency becomes less critical. Also, there are some possible problems with high ERs. Lasers don't like being asked to turn off and then quickly turn back on. An almost "off" condition is needed for a high ER. As the laser turns back on, it can experience some significant shifts in output wavelength. As fiber does not have a constant propagation velocity for all wavelengths, transmitted pulse can be dispersed in time. Almost turning the laser off can induce a jitter mechanism, as the relative time at which the laser turns on may be inconsistent. If ER specifications are relaxed, reducing some of the issues just discussed, it is still important to maintain a good separation between 1's and 0's at the receiver. As system power budgets were designed for the early optical Ethernet and Fibre Channel systems, this separation in signal levels, and not extinction ratio, became the important parameter to measure. This separation is referred to as "optical modulation amplitude" (OMA), which is computed by taking the difference in the amplitude of the 1 and 0 levels.

Note that one signal with high extinction ratio may have the same OMA as another signal with low extinction ratio. Similarly, a signal with large OMA may have the same extinction ratio as a signal with low OMA. The parameter that links OMA and extinction ratio is the average power of the signal. For the two signals with common OMA, the higher extinction ratio signal will have a lower average power. This makes sense, as high extinction ratio implies that for a given modulation content, less transmitter power is used. The relationship between ER, OMA, and average power is:

$$\text{OMA} = \frac{\text{ER-1}}{\text{ER+1}} \cdot \text{AveragePower}$$

OMA should easily fall out of the computations that the oscilloscope makes for extinction ratio. However, communications standards, specifically 802.3 Ethernet standards have put an important twist in the measurement of OMA. When setting the system level power budgets, OMA is considered without impairments like ISI and waveform transients, which are handled elsewhere. Thus an ideal OMA value should be measured. This is achieved by having the transmitter put out a square wave pattern, typically five logic ones followed by five logic zeros. The "one" amplitude is taken from a histogram across the middle bit of the run of ones and the "zero" amplitude is taken across the middle bit of the run of zeros. By taking data that is far from any bit edges, the ideal OMA is obtained.

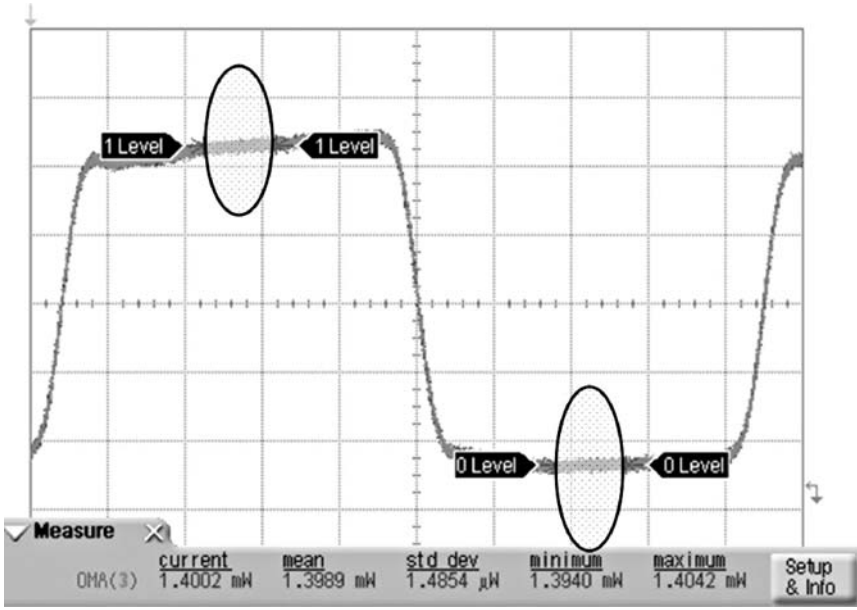


Figure 14.17 OMA is measured on a square wave pattern. The one level is obtained from samples from the central 20% of the sequence of ones and the zero level is obtained from samples from the central 20% of the sequence of zeroes.

14.13. GOLDEN PLL TRIGGERING FOR TRANSMITTER TESTING

As speeds increased up to and beyond 1 Gb/s, the timing stability of signals became more difficult to manage. Timing jitter, observed when signal edges are not consistently located in time, leads to decision circuits not making their decisions in the center of the bit. As edges drift towards what should be the ideal decision time, and the receiver begins to make decisions at a point along those edges instead of at the center of the bit, the bit error ratio (BER) is degraded. To avoid this, transmitter jitter must be controlled to manageable levels, and once again the capabilities of the system receiver dictate what this level should be.

Receivers require a clock signal to time their decision process. Many receivers derive this clock directly from the incoming data stream through some form of clock extraction circuit. The clock extraction process provides some tolerance to transmitter jitter, because the receiver clock extraction circuitry can track and follow jitter in the incoming data stream, as long as the jitter is not too fast or too large. Typically, if the rate of the jitter is within the loop bandwidth of the clock extraction circuit, the receiver will tolerate it. There is no exact rule for

setting the receiver loop bandwidth. However, values on the order of 1 to 5 MHz are common. This then sets a rough boundary for what would be considered low frequency jitter (well below the loop bandwidth value) and what would be considered high frequency jitter (well above the loop bandwidth).

If receivers are tolerant to lower rate jitter, it does not make sense to reject transmitters that have jitter at these low rates. In recent years, communications standards have been designed to account for this issue. The oscilloscope used to measure jitter is specified to have a high-pass jitter function so that test results are not impacted by the presence of low frequency jitter. The easiest way to produce a jitter high-pass function is to derive the oscilloscope triggering from the observed waveform, in a manner similar to how the receiver derives its clock. This is often referred to as "Golden PLL" testing. A clock recovery circuit with a specific loop bandwidth is built into the oscilloscope, allowing it to generate timing (or triggering) signals similar to those generated by a system receiver.

Figure 14.18 shows that when the test system employs the correct Golden PLL bandwidth, the test system mimics the system level receiver and eliminates low frequency jitter, as low frequency jitter is common to both the oscilloscope trigger and the observed signal. Both measurements are of the same signal, but the lower waveform, using the clock extraction circuit with the correct loop bandwidth, provides a better assessment of the waveform from the perspective of the receiver it will be paired with.

14.14. JITTER ANALYSIS

As speeds increase, it becomes more difficult to maintain the horizontal opening of the eye diagram. The result is similar to vertical eye closure. That is, receivers will have a more difficult time determining the logic level of bits. BER is then degraded. While vertical eye opening is reduced through noise, waveform distortion, and attenuation, horizontal eye opening is greatly affected by the timing stability of the transmitted bits. Timing instability is commonly referred to as jitter.

Jitter, in its most basic definition, is the deviation of the edges of a data signal from their ideal positions in time. To understand why jitter can degrade BER, consider that a data receiver must set an optimum signal level and an optimum time to determine if a received bit is at a high or low logic level. The decision threshold is typically the center *amplitude* of the bitstream. The decision process is degraded when the waveform has excess noise or distortion or is small in amplitude. The decision is ideally made in the *time* center of the bit. The decision process is degraded if the timing of the bits is inconsistent. A jitter free signal would be one that has unvarying bit periods and edges that occur exactly at the expected times. This allows the receiver to make its logic decision far away (in time) from where bits transition from low to high levels or vice versa. If the

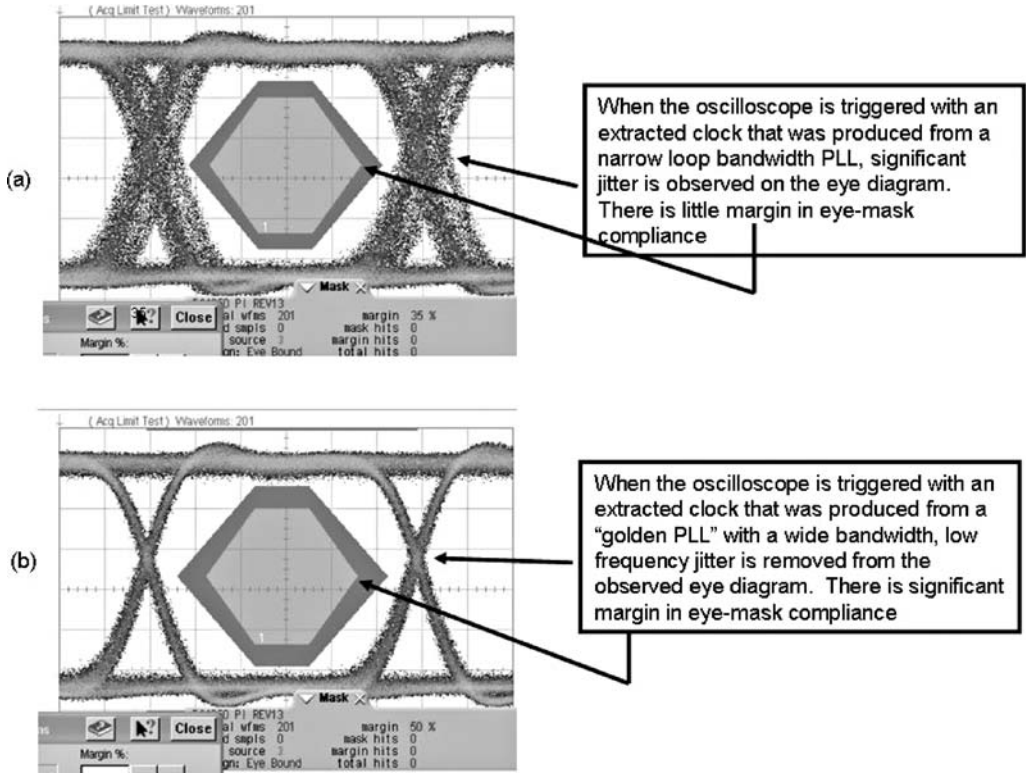

(a)
When the oscilloscope is triggered with an extracted clock that was produced from a narrow loop bandwidth PLL, significant jitter is observed on the eye diagram. There is little margin in eye-mask compliance
(b)
When the oscilloscope is triggered with an extracted clock that was produced from a "golden PLL" with a wide bandwidth, low frequency jitter is removed from the observed eye diagram. There is significant margin in eye-mask compliance
Close
Mask

Figure 14.18 Waveform test with a near jitter free trigger (a) and a Golden PLL trigger (b). The observed jitter can be significantly reduced when the timing reference (trigger) jitter is common with the jitter on the signal being observed.

receiver makes its decision near the transition (edge), the probability of the receiver making a mistake increases.

One can visualize jitter through the eye diagram. See Figure 14.19. The eye diagram is a composite of all the bits from a data stream overlaid on a common time base. If the timebase is defined by a jitter free clock signal at the nominal data rate, then some bits will be slid to the left (early bits) or to the right (late bits) when there is jitter. This results in horizontal closure of the eye diagram. In real world systems, the most extreme displaced edges (those closest to the center of the eye, assuming the jitter magnitude is not extreme) occur least frequently. More frequently, edges will be located close to their ideal positions. See Figure 14.19. If this eye diagram were jitter free, all the one to zero transitions would be tightly grouped together, as would all the zero to one transitions. But in this case, there is a noticeable difference between the earliest and latest bit edges.

As data rates increase, jitter problems tend to become more difficult. What might have been considered a small and tolerable time deviation at a lower data rate appears to be large and intolerable at high data rates since the bit period has become proportionally smaller. Measurement tools to accurately quantify jitter are required. In addition to quantifying jitter, methods that provide insight into the nature of jitter and its root causes become more important. System bit-error-ratio specifications at levels of 1E-12 and lower require the ability to quantify

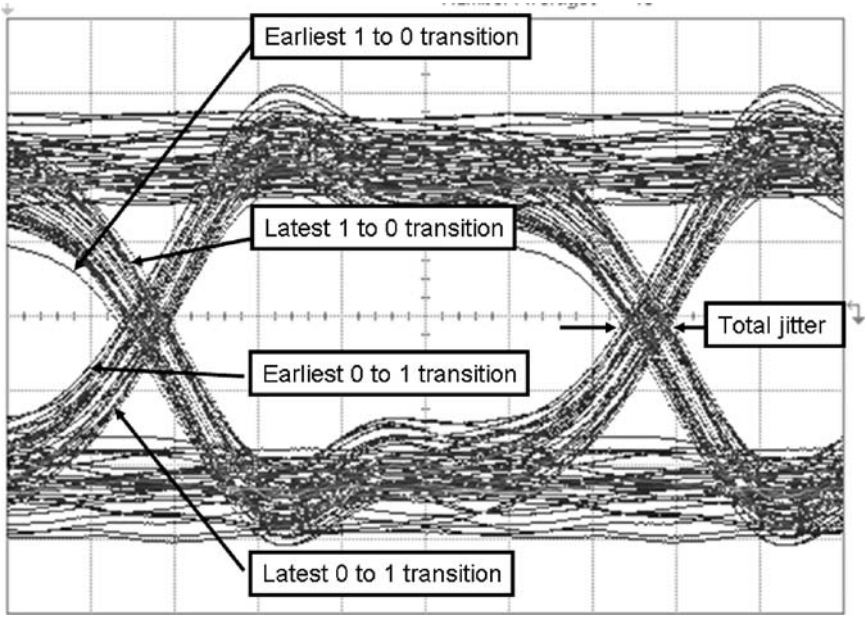


Figure 14.19 Eye diagram for a transmitter with significant jitter.

jitter at extremely low levels and probabilities. The analysis of jitter can be enhanced through segregating jitter into classifications aligned with fundamental root causes. The first two groupings are for jitter due to random mechanisms (RJ) and jitter due to deterministic mechanisms (DJ). RJ is the jitter due to stochastic processes such as those that cause random fluctuations in oscillator/clock frequencies, which in turn cause data rates to fluctuate proportionally. Conversely, DJ is jitter that is due to predictable mechanisms. The DJ can be further classified into two subclasses. Some DJ will be due to how a transmitter performs (i.e., how jitter is produced) for different data patterns. This is data dependent jitter (DDJ). Other elements of DJ will be deterministic, but uncorrelated to the data pattern. For example, switching power supply noise, which is periodic, but unrelated to the data pattern, can cause data clocks to deviate in frequency and lead to jitter of the data.

Optical waveforms are most often analyzed with sampling oscilloscopes, as this class of oscilloscopes have the widest bandwidths, a requirement for accurate analysis of very high-speed digital transmitter. However, the wide bandwidth has historically been accompanied by a relatively low data acquisition rate. This has significantly limited the capability of the sampling oscilloscope to a coarse view of the overall jitter of a signal.

New developments in sampling oscilloscopes have radically changed the jitter analysis capabilities available for very high-speed transmitters. Jitter can now be decomposed into its random and deterministic components including periodic, inter-symbol interference, subrate, and duty cycle distortion jitter classifications. The measurement process is incredibly fast due to advanced triggering capabilities that allow efficient analysis of long data patterns, which is essential for practical analysis of data dependent jitter.

Figure 14.20 shows the jitter analysis results for a transmitter operating at 4.25 Gb/s. The jitter of this signal has been decomposed into its constituent components. The individual jitter mechanisms can then be combined to predict what the aggregate jitter magnitude is to a 10E-12 probability, which can then be used to design system level jitter budgets that operate at low BER's.

Most communications systems are built around an industry standard. The standard in turn will have the minimal specification for the components that make up the system. A battery of tests, many based upon analysis of the eye–diagram, are used to determine the viability of the transmitter to work in a complete system. As transmission speeds increase and bit periods shrink, in-depth analysis of timing stability (jitter) has gained importance in transmitter verification.

14.15. RECEIVER TEST

The receiver must be capable of taking the signal from the output of the fiber and determining the logic level of each bit. Since signals from the transmitter will

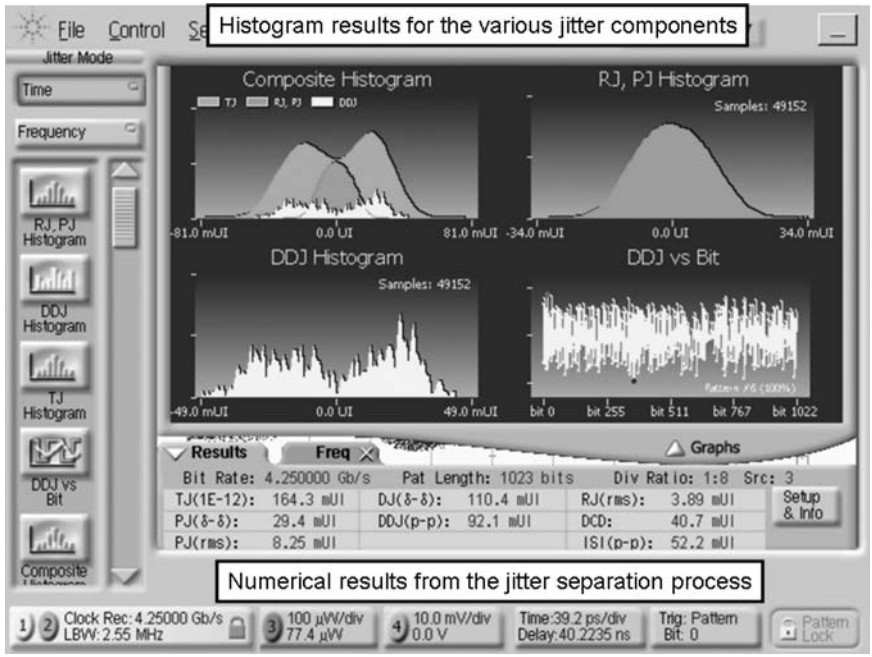


Figure 14.20 Jitter separation analysis of a 4.25 Gb/s signal.

not be perfect, and as the signal traverses the fiber it will be further degraded through attenuation and dispersion, verifying the performance of a receiver focuses on determining how often the receiver makes mistakes (calls a received "1" a "0" or vice versa). Thus the receiver test system is one that provides a test signal that replicates the types of signals that the receiver might encounter in an actual system. The test system must also be able to verify whether the receiver made any mistakes. This measurement is achieved with a bit-error-ratio tester (BERT). The BERT consists of a pattern generator and an error detector.

14.16. PATTERN GENERATOR

A pattern generator produces a known sequence of binary data. The most common sequence is a pseudo-random binary sequence or PRBS. The PRBS pattern is easy to generate and has the following properties:

- A systematic, repeatable method to produce data patterns that approach characteristics similar to truly random data
- Pattern lengths are 2^N-1
  - Common values for N: 7, 10, 15, 23 and 31

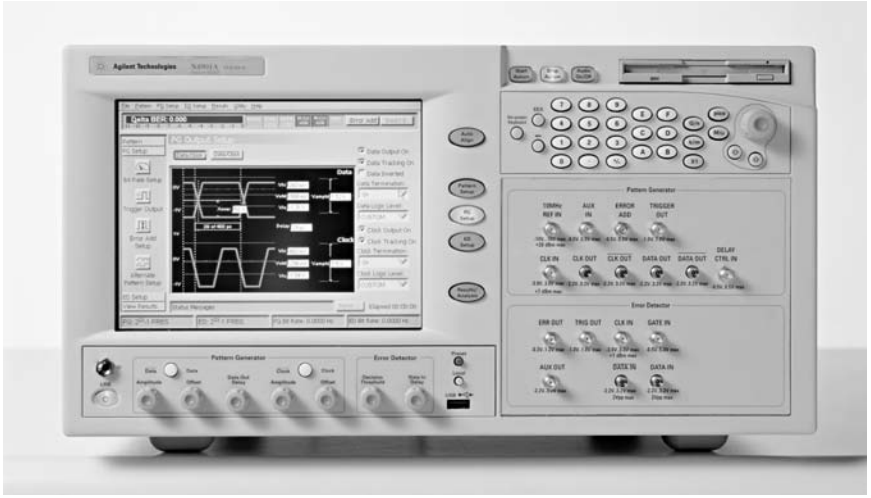


Figure 14.21 Common BERT includes both a pattern generator and an error detector section.

- The 2^N-1 will have all possible combinations of N bits (except N consecutive 0's)
- Generated through N cascaded shift registers with a feedback tap
- They simulate "random data"
- The data sequence is deterministic
- The pattern repeats and can be predicted.
- Easy to generate and measure at high speed
- When you compare the sequence with itself, you get 50% errors in all positions except exact pattern alignment.
- Easy to vary the ones (mark) density
- You can put the pattern out of balance in a randomly distributed way.
- Not harmonically related to the data rate (2n-1 only)
- Broad spectral content allows use as a noise source

Modern BERTs also have pattern generators that can "precisely" corrupt the signal they produce in order to simulate actual signals a receiver would encounter in a real system. This intentional signal corruption is commonly referred to as signal "stress". A stressed signal will typically have timing jitter, bandwidth limiting (inter-symbol interference), and attenuation.

14.17. THE ERROR DETECTOR

The error detector is used to determine if the data received matches the transmitted pattern. The error detector will receive the output of the receiver under

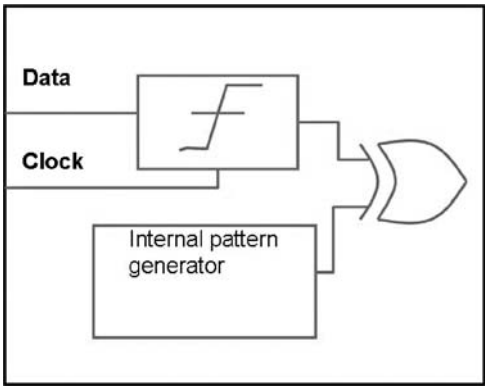


Figure 14.22 The error detector consists of a reference pattern generator an a exclusive OR comparator.

test. In its most basic building blocks the error detector consists of an internal pattern generator and an exclusive OR logic gate. The error detectors internal pattern generator will produce a reference pattern identical to that presented to the receiver. The reference pattern is synchronized to the test device output and then compared bit for bit with the exclusive OR gate. Any disagreement as counted as a bit error. Determining the BER is then a simple matter of counting the total bits received and the number that were received in error.

A common receiver test is one of sensitivity. Most receivers will operate essentially error free if the signal they take in has a large power level. As power is decreased, eventually the receiver will start to make mistakes. A useful receiver measurement is BER as a function of input power. An attenuator is placed between the transmitter (fed by the pattern generator) and the receiver. The BER is measured at several attenuator/power level settings.

This leads to the concept of "power penalty" measuremnts. Rather than testing BER as a function of received power, BER can be tested as a function of other impairments. Examples of impairments include fiber dispersion or distance, types of signal regeneration, clock recovery schemes and so on. The test essentially determines "How much does the signal strength need to be increased to achieve the same BER system had when the signal impairment was not present?" The basic process is as follows:

- Example:
  - Measure system BER without impairment
  - Set system attenuation for desirable BER (1E-10)
    - Measure power at receiver (–22 dBm)
  - Add dispersion (e.g., spool of fiber)

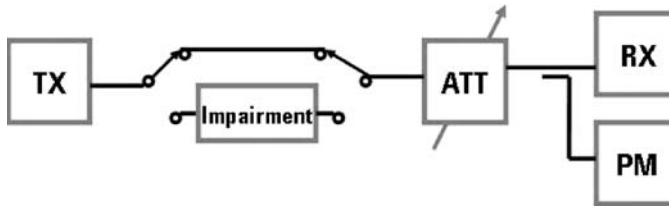


Figure 14.23 Configuration for a power penalty measurement.

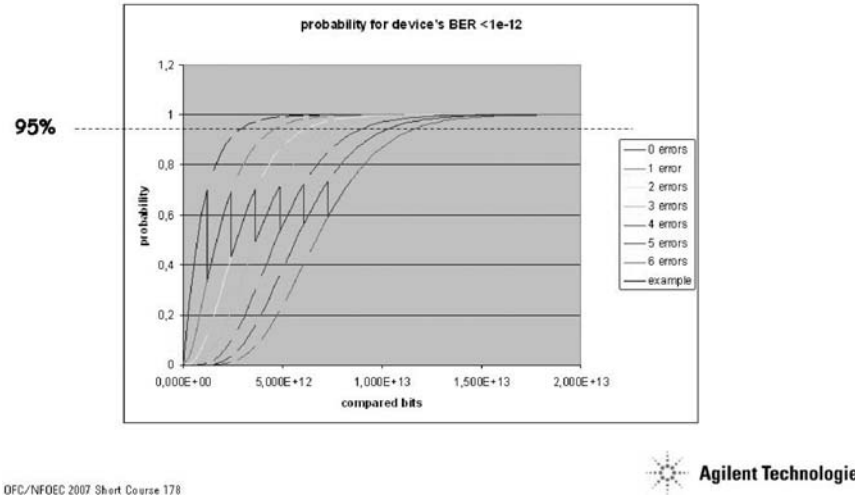


Figure 14.24 Confidence levels for various numbers of errors.

- Reduce attenuation to return to 1E-10 BER
  - Measure power at receiver (–18 dBm)
- Difference in power levels is dispersion power penalty (4 dB)

Since BER measurements are used to characterize the effects of impairments such as noise, they require analysis from a statistical perspective. Specifically, one must consider how many errors must be detected to render a statistically valid measurement. Ideally, a minimum of 1000 errors should be measured to yield a high confidence level in the accuracy of the BER measurements, although good measurements with as few as 100 errors can be achieved. This leads to one of the primary difficulties encountered when performing BER tests. If a BER of 1E-10 is required, and at least 1000 errors are to be detected, then at least 10 terabits must be transmitted to during the measurement. If the transmission rate is

2488 Mb/s, the time required for the measurement will be over an hour. If the data rate is 155 Mb/s, the test time would be almost 18 hours!

Is there a way to make the measurement faster and still be confident in the result? What happens if no errors are measured? How confident can you be in the BER performance? The following formula can help:

$$C = 1 - e^{-nb}$$

Where

C = degree of confidence, (0.95 = 95% confidence)
n = number of bits examined with no error detected
b = desired residual BER
Example: With no errors over 3E12 bits, 95% confidence BER is better than 1E-12

The above formula works when no errors are measured. What happens when one or more errors occur. Rather than provide several formulas, the following graph can be used to determine confidence level versus number of bits received and the number of bits received in error. Thus a BER result can be obtained with good confidence even when very few errors are observed.

This page intentionally left blank

15

Optical Wavelength Division Multiplexing for Data Communication Networks

Klaus Grobe
ADVA Optical Networking, 82152 Martinsried, Germany

15.1. BASICS OF WDM SYSTEMS

15.1.1. CWDM and DWDM

Wavelength-division multiplexing (WDM) enables multiple-shift usage of transmission fibers by coupling several wavelengths into the fibers through appropriate optical filters.

Two WDM flavors are standardized; dense WDM (DWDM) according to ITU-T G.694.1, and coarse WDM (CWDM) according to G.694.2. For DWDM, a channel grid of 12.5/25/50/100 GHz has been defined, with 200 GHz, 400 GHz, and so on, also possible. For example, the 100-GHz grid starts at 195.90 THz and ends at 184.50 THz (~1530 . . . ~1625 nm), defining 229 equidistant frequencies. In long-haul (LH) and ultra LH (ULH) areas, 50 GHz are often used; for metro and regional systems, 100 GHz or 200 GHz are commonly used due to cost advantages. Also, 25 GHz and 12.5 GHz are investigated for ultra-dense WDM systems. Figure 15.1 (top) shows an example of the frequency allocation of a 200-GHz system.

CWDM is equidistant in the wavelength domain (as compared to DWDM, which is equidistant in frequency), the channel spacing being 20 nm. So far, 18-CWDM channels are defined, ranging from 1270 to 1610 nm. Most commercial systems offer 8 CWDM channels (1470 to 1610 nm), the reason being the limited roll-out of metro fibers with reduced OH-absorption according to ITU-T G.652C/D (e.g., Lucent AllWave®). Figure 15.1 (bottom) shows these 8 CWDM channels relative to the DWDM channels.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

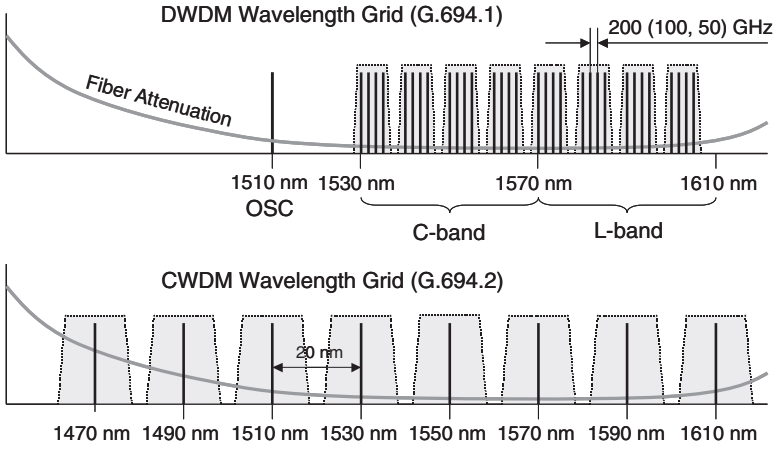


Figure 15.1 DWDM channels (top) and CWDM channels (bottom).

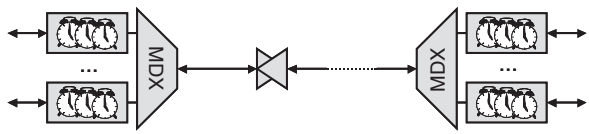


Figure 15.2 Basic WDM system. Also shown is a (bidirectional) inline amplifier.

WDM wavelengths are provided by transponders or colored interfaces in the client equipment. Channel cards with two or more client interfaces and time-division multiplexing (TDM) are referred to as muxponders. Trans- and muxponders provide ITU wavelengths, power-level adaptation, pulse shaping, and optional de-jittering (2R or 3R regeneration). They also act as demarcation between client and transport system. In order to support various client signals, they can have multiple clocks for 3R. A basic WDM system comprises several transponders (muxponders, colored interfaces) and filters for multiplexing/demultiplexing (MDX) (Fig. 15.2).

A 3R transponder is shown in Fig. 15.3. The diagram shows an external optical modulator (EOM), which is typically used for high (line) bit rates. Also shown is the clock recovery and a feedback loop in the transmit part, which is used for wave locking. Active wave locking is necessary for grids at 100 GHz and below in order to control the filter passbands.

The MDX filters can provide access to all wavelength channels where appropriate. This is the case in optical termination multiplexers (OTMs) where all

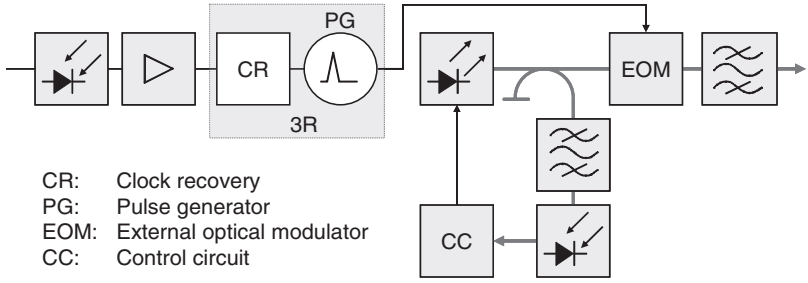


Figure 15.3 3R transponders with wave locker.

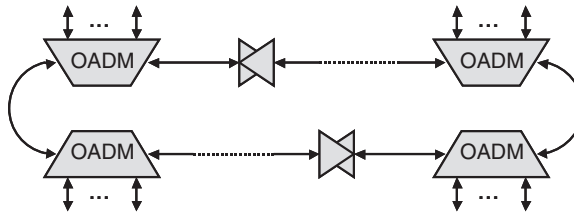


Figure 15.4 WDM ring consisting of OADMs and ILAs.

channels are terminated. Examples are all point-to-point (P2P) links. Between two OTMs, access to only a subset of channels may be required. This leads to the concept of the optical add/drop multiplexer (OADM). OADMs provide access to a subset of WDM channels; the other channels are passively passed through. This avoids unnecessary and costly termination of WDM channels with transponders or colored interfaces [1]. Similar to SONET/SDH, OADMs can be used to build WDM rings. These rings support various optical (WDM) protection schemes, and they can make efficient use of fiber infrastructure when trying to connect various sites via diverse routes. A simplified WDM ring with 4 OADMs and additional inline amplifiers (ILA) is shown in Fig. 15.4.

OADMs often make use of two-stage filters. The first stage (band or group splitters) splits the wavelength spectrum into a certain number of groups or subbands. These groups can be added, dropped, or passively passed through. Groups that are terminated use a second filter stage to provide access to individual wavelengths. This leads to filter designs that allow capacity upgrades, which do not affect service.

Different technologies are used for WDM filters, depending on flexibility, performance, and cost requirements. Modular DWDM optics and almost all CWDM optics make use of thin-film filters (TFFs). In a TFF, thin dielectric,

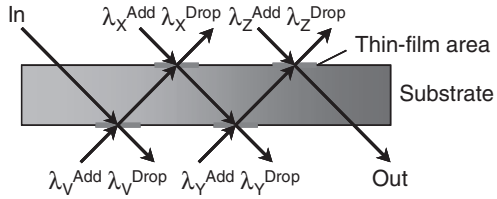


Figure 15.5 Thin-film filter.

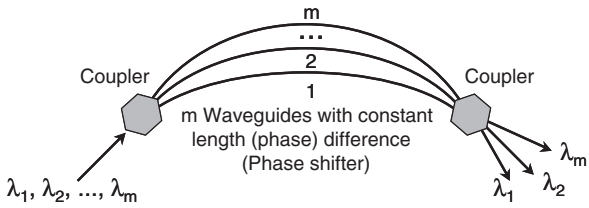


Figure 15.6 Arrayed waveguide WDM filter.

wavelength-selective layers are applied to a substrate (see Fig. 15.5). Light can traverse the substrate through the effect of total internal reflection, or it is coupled/decoupled in the corresponding thin-film areas. Thus, wavelength decomposition is provided through spatial decomposition.

Alternatives to TFF include fiber bragg gratings (FBGs) and arrayed waveguide gratings (AWGs). FBGs are periodic density variations inside fibers (see Chapter 2). These density variations are tuned to individual wavelengths; the corresponding wavelengths are then reflected by the respective FBG, rather than being transmitted. FBGs can be combined with fiber interferometers or circulators in order to separate the transmission directions.

AWGs are compact, single-stage filters that provide access to a high number of wavelengths (typically 40). They consist of couplers and waveguides that are integrated into a substrate. The waveguides have different lengths that lead to constructive or destructive interferences in the output coupler, and hence multiplexing/demultiplexing (see Fig. 15.6).

AWGs are used in static P2P configurations, for example, high-capacity storage applications, or flexible, single-channel OADMs (so-called ROADMs; refer to Chapter 15.5.1).

Table 15.1 lists relevant parameters of the filter technologies.

Both the multiplexing/demultiplexing and modulation techniques have been improved in recent years. Early LH WDM systems were capable of transmitting 8×2.5 Gb/s over distances of 600/640 km (according to ITU-T G.692). Latest-generation metro/regional WDM systems cover the same distances with up to

Table 15.1
Parameters of WDM Filter Technologies.

	AWG	TFF	FBG
Channel spacing [GHz]	200/100/50	400/200/100	200/100/50
Insertion loss	Moderate	Low	High
Bandwidth @ –0.5 dB	Narrow	Broad	Broad
Bandwidth @ –25 dB	Broad	Broad	Narrow
Sidelobe suppression	Moderate	High	High
Passband ripple	Moderate	High	Low
Crosstalk, neighbor	Low	Low	Very low
Crosstalk, nonneighbor	Low	Very low	Very low
Polarization dependence	High	Low	Low

80×10 Gb/s or even 40×40 Gb/s, whereas ULH systems cover distances of 2000 to 3000 km with up to 160×10 Gb/s.

15.1.2. WDM and TDM

WDM systems can transport high numbers of different services simultaneously. However, cost-effective transport is possible only if the wavelengths are utilized with high-service bandwidths. Transport of services with 100 to 1000 Mb/s or even 2.5 Gb/s per wavelength is no longer economic. Hence, services at low bit rates must be aggregated onto wavelengths with 10 Gb/s or 40/100 Gb/s in the future.

WDM systems can be combined with synchronous multiplexing as known from SONET/SDH. (Alternatively, SONET/SDH ADMs can make use of colored WDM interfaces.) Similar aggregation can be provided for Ethernet services (i.e., WDM systems with Layer-2 (L2) functionality, or switches with colored interfaces). In addition, asynchronous TDM functionality for low-rate data and storage services has been developed. Early variants made use of proprietary, per-bit or per-byte asynchronous interleaving. This way, muxponders for $8 \times$ ESCON or $2 \times$ FC/GbE and 2.5 Gb/s line rate were provided, leading to WDM systems that supported >500 ESCON channels. Sometimes an ESCON or Fibre Channel/FICON switch employing TDM would be used in conjunction with a dedicated WDM platform, such that intra-switch links could be run over WDM. Early examples include the FICON Bridge feature on some ESCON Directors (see Chapter 21). Straight per-bit or per-byte interleaving rather than container-based multiplexing (which was also developed as proprietary TDM) is relevant for many data and storage applications in order to guarantee low latency. The TDM schemes can be combined with SONET/SDH framing (to provide monitoring and interworking capabilities). The resulting network structure is shown in Fig. 15.7.

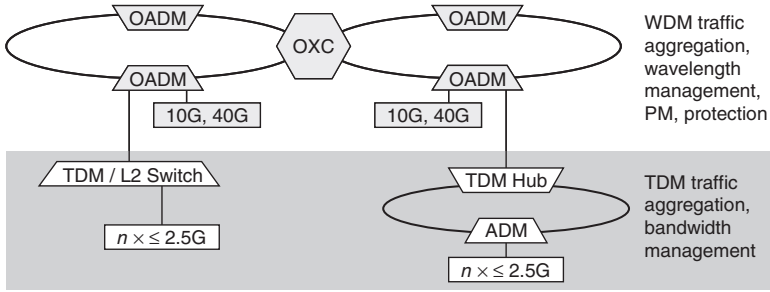


Figure 15.7 Combination of WDM and TDM in a metro/regional network.

Today, TDM is based mainly on standards, that is, Generic Framing Procedure (ITU-T G.7041) for data and storage services (see Chapter 19), and G.709 for multiplexing in optical transport networks (OTNs).

15.1.3. Metro and Regional WDM Networks

WDM systems are the transport-layer basis of metro and regional networks (as well as of LH/ULH networks). They support legacy SONET/SDH networks as well as all data and storage appications. Beyond transport, WDM can provide performance monitoring (PM) and further functionalities such as optical protection [2].

Metro WDM networks often consist of WDM rings (both DWDM and CWDM). Rings lead to the most economic solution for access via diverse paths from network operators' points of presence (PoP) to multiple sites. DWDM rings typically have 32 to 40 wavelengths, ready for 10 Gb/s (10G) or even 40G per-channel bit rates. This leads to maximum aggregated capacities of up to 400/1600 Gb/s. CWDM rings typically have 8 (or 16) wavelengths that support channel bit rates of 2G5. A subset of the wavelengths may also support higher rates of 4G or 10G, respectively [3]. CWDM rings almost always provide static single-channel add/drop, whereas DWDM rings make use of static or flexible banded (group) or unbanded (single-channel) add/drop. This is complemented by compensation of chromatic dispersion (CD) through dispersion-compensating modules (DCM) and optical supervisory channels (OSC).

For regional networks (up to 1000 km), WDM systems are often configured as linear add/drop (LAD) systems. These systems require dispersion compensators (DCM) and transponders with sufficient CD allowance. They also require amplifiers with reasonably low noise figures and sufficient gain flattening and an OSC in order to provide management channels for the ILAs. A regional LAD system is shown in Fig. 15.8.

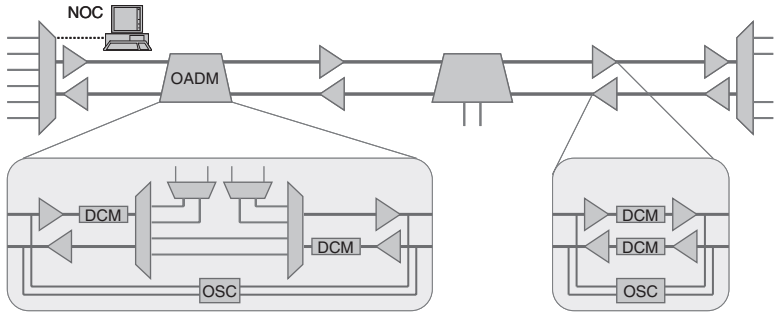


Figure 15.8 Regional LAD-WDM system.

Compared to other transport technologies, WDM has very few disadvantages. Depending on the configuration and requirements, there may be significant up-front invest for WDM systems that support high channel loads, but start only with a low number of WDM channels. This can be reduced by WDM systems that provide high levels of modularity, flexibility, granularity, and hence pay-as-you-grow.

15.1.4. Protection and Restoration

Regional and in particular metro networks require resilience mechanisms in order to provide certain levels of path availabilities for sensitive appications. Here, path availability is the total end-to-end (E2E) availability of an optical path, that is, between client interfaces (e.g., FC directors).

Resilience has to be split into protection and restoration [4]. For protection, the required resilience capacity is preassigned or at least (shared protection) pre-calculated. For restoration, the restored path is calculated only after a failure. Hence, protection in most cases is simpler and faster. Metro WDM systems almost exclusively offer protection. Restoration is discussed in the context of automatic switched optical networks (ASON) or generalized multi-protocol label switching (GMPLS).

15.1.4.1. WDM Protection Mechanisms

For IP/MPLS-over-WDM and further applications, a flexible photonic layer providing intelligent resiliency is required. For backbone networks with (partly) meshed traffic, restoration is appropriate. This will be enabled by means of a distributed (GMPLS) control plane. In metro areas, typically rings are deployed. Here, restoration is replaced by (WDM) protection.

Overviews on optical protection can be found in [4–6], for example. Various solutions for WDM protection exist:

- Ring protection vs. span/line switching
- Optical channels (OCh) vs. optical multiplex section (OMS) vs. optical wavelength groups
- Spare fibers vs. spare wavelengths/spare subsets of wavelengths
- Dedicated protection vs. shared protection
- Unidirectional rings vs. bidirectional rings

The last two items are not independent from each other since shared protection rings always have to be bidirectional rings.

Dedicated protection is also known as 1 + 1 or 1:1 protection, and 100% redundancy (transmission capacity, fiber, equipment) is assigned to the working capacity. Dedicated protection rings usually are 2-fiber path-switched rings and can be divided into unidirectional and bidirectional rings. Unidirectional rings use one fiber and one direction for the working paths and the second fiber and the counter direction for protection. Bidirectional rings always use both fibers and the same spans for both directions of the paths. Dedicated protection rings are the appropriate choice for hubbed traffic patterns. In this case, the maximum number of routable connections P between a hub and decentralized OADMs equals the number of wavelengths W:

$$P = W \tag{15.1}$$

With shared protection, several working signals share a common protection capacity. It is also named 1:N protection, based on 4-fiber or 2-fiber rings. 2-Fiber rings then require single-fiber working or special assignment of different wavelengths to the counterpropagating signals. In order for different signals to be able to share common protection capacities, meshed traffic is required. In this case, shared protection can save substantial amounts of CapEx or increase the total ring capacity. If only site-to-adjacent-site (SAS) traffic is assumed, the maximum number of routable paths becomes

$$P = W \cdot N \tag{15.2}$$

where N is the number of nodes. In Eqs. (15.1) and (15.2), W has to be replaced by the number of wavelength groups G for some metro WDM installations since 2-stage filters with group splitters providing the add/drop are used.

Optical protection rings are sometimes referred to as UPSR, BPSR, or BLSR. However, *paths* (i.e., wavelengths) in a WDM system usually are *lines* in a SONET/SDH system, which can lead to confusion for the classification of WDM protection schemes. In addition, WDM systems offer a new level of granularity, *groups of wavelengths*. Table 15.2 lists two sets of terms for optical protection rings.

The first set is based on SDH, and the second on SONET terminology. A capital "O" has been prefixed to all the terms in order to differentiate them from SONET/

Table 15.2
Optical Protection Rings.

	Dedicated	Shared
OMS (line)	—	OMS-SPRing or O-BLSR
Wavelengths Group	Not standardized	—
OCh, path	OCh-DPRing or O-UPSR/O-BPSR	OCh-SPRing or O-BPSR

SDH wording. Concerning the optical wavelengths group layer, no standardized wording exists. Some products employ nonstandardized solutions, such as optical trunk switches, which may be configured to switch between a working and protected WDM path but not to switch individual wavelengths. Switching speeds may be slower than the 50-ms SONET/SDH standard (up to 100 ms or more). Variants on this approach use splitters or Y-cables to provide redundancy for client adapters.

OCh-DPRings/O-BPSRs are the majority of WDM protection rings (>90%). Client signals are fed into both ring directions (east, west) using duplicated transponders or switches. Switching is accomplished in the receive ends, individually per channel, and without intervention of a centralized management system. This leads to reliable and fast switching. A disadvantage is the hardware effort, in particular if transponders (and filters etc.) are duplicated. A 4-node path-switched ring is shown in Fig. 15.9.

A 4-fiber O-BLSR shared-protection ring has also been reported [7]. Here, a second fiber pair is used for protection. For meshed traffic, this ring can support higher payload capacities with the same number of wavelengths, but provisioning as a line-switched ring is inflexible and complex. In the next two sections, two fairly new ring protection concepts are discussed [3, 8].

15.1.4.2. Concept of an OCh-SPRing

The OCh-SPRing discussed here provides shared protection, which is provisioned on a per-wavelength basis. One-half of the wavelengths is used for 1 : N protection. In case of a failure, they are assigned to the affected working wavelengths. As shown in Fig. 15.10, switching is accomplished by the OADMs that terminate the affected span.

In case of a fiber failure, traffic running through the affected span is able to allocate the protection wavelengths. For meshed traffic, working wavelengths can be reused, which leads to the sharing effect for the protection wavelengths and consequently higher total ring capacity. An advantage over the O-BLSR [7] is the per-channel provisioning, which also allows combinations with other protection schemes (client-layer protection, 1 + 1 dedicated protection, unprotected).

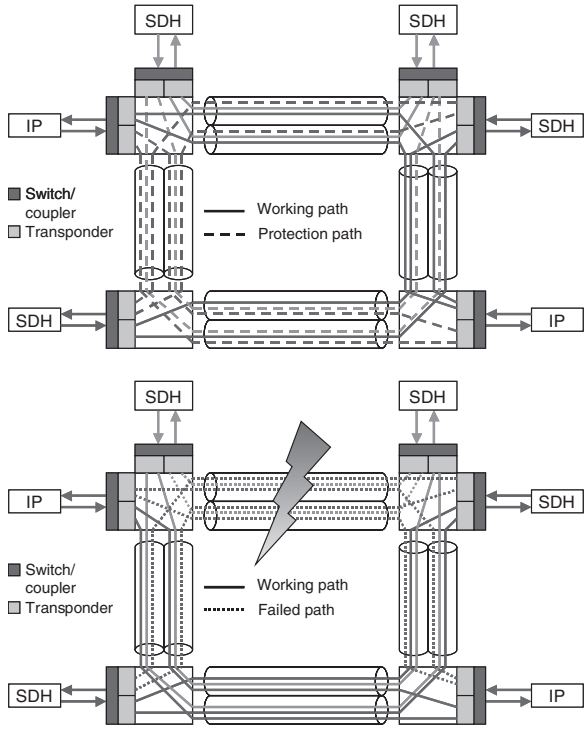


Figure 15.9 Optical channel dedicated protection ring (OCh-DPRing).

OCh-SPRings can lead to an extension of the maximum lengths of the protection paths that is typical for shared-protection rings. The longest possible protection path length is given by:

$$L_{prot} = 2 \cdot C_{ring} - 3 \cdot L_{span} \tag{15.3}$$

L_{prot} is the maximum protection link length, C_{ring} is the ring circumference, and L_{span} is the (mean) length of a single span, respectively. This effect is shown in Fig. 15.11.

The same effect can be found for SONET/SDH *line-switched* rings and the O-BLSR. Standard O-XPSR lead to shorter maximum link lengths of only $C_{ring} - L_{span}$. This effect has to be considered for delay-sensitive applications, and for link engineering (compensation of CD). For meshed traffic and in particular SAS traffic (which is typical for SONET/SDH as payload!), OCh-SPRings can lead to a significant increase of total ring capacity and cost reduction. For hubbed traffic, there is no advantage over OCh-DPRings.

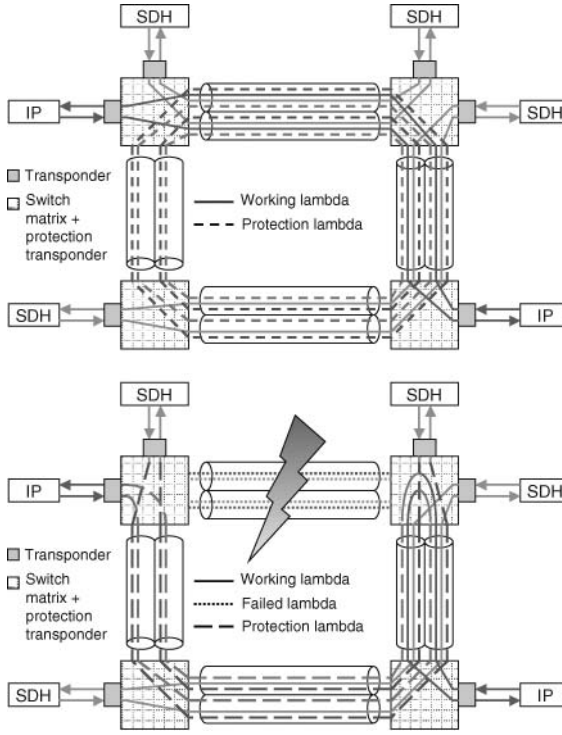


Figure 15.10 Switching event in an OCh-SPRing.

Since the total ring capacity is increased for a given set of hardware and fibers, the corresponding cost per transported bit/s is decreased. Although more hardware effort is required as compared to an OCh-DPRing, this leads to cost savings for OCh-SPRings, depending on the traffic pattern.

15.1.4.3. Concept of a Dedicated Group Protection Ring

The Group Protection Ring (GPR) combines the simplicity of dedicated protection with the cost savings known from shared protection or line switching. It is based on group-splitting filters (GS) and subsequent channel mux/demuxes. Here, the groups are split by means of passive 3 dB couplers and then fed into both ring directions (east, west). At the receive end, both group signals are fed into a protection switch (PS), which selects the east or west direction upon loss of light (LoL). The PS can switch per individual group. A simple 3-node GPR is shown in Fig. 15.12. Similar to all DPRings, the switching is accomplished in the

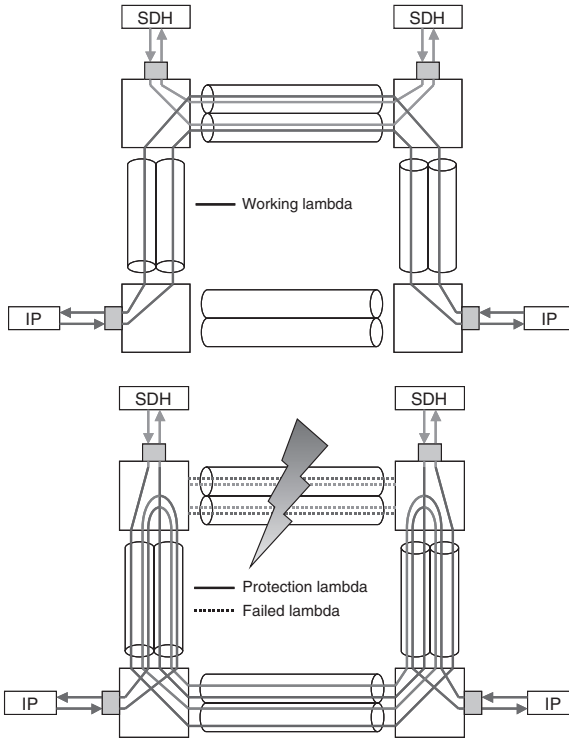


Figure 15.11 Extension of protection path length in a SPRing.

nodes that terminate the affected signals. This leads to the maximum protection link extension of $C_{\text{ring}} - L_{\text{span}}$, as already noted.

The groups can switch autonomously and independently from each other. Hence, fast switching within 60 ms is possible, and combinations with client-layer protection and unprotected traffic are possible. Also, no complex signaling in the ring is necessary.

The GPR can lead to substantial cost savings over an OCh-DPRing in the range of 25–40%. Total resulting path availability is decreased (from ~99.999% for an OCh-DPRing) to ~99.995% due to the lack of transponder protection. This is tolerable for many metro WDM applications.

15.1.5. Optical Switching

Optical switching is necessary for protection/restoration switching or flexibility, which is provided in ROADMs. The corresponding switching technologies

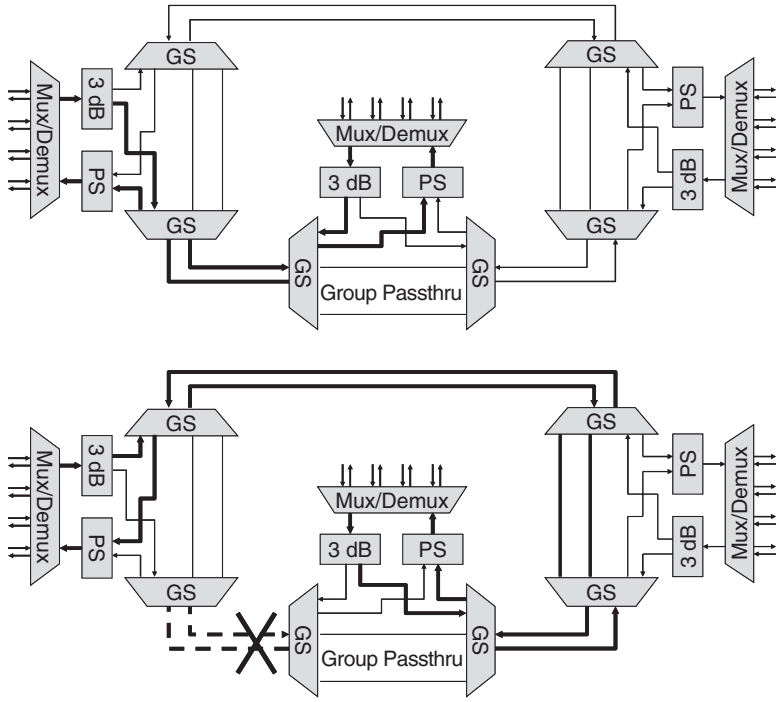


Figure 15.12 Protection event in a GPR.

must have low insertion loss, high isolation, and low backreflection. They must also support high availability and scalability, compact design, fast switching time, and low power consumption.

Relevant switching technologies include micro-electro-mechanical switching (MEMS), liquid crystal technology (LCT), bubbles, thermo-optics, acousto-optics, and holograms. Certain additional techniques like liquid gratings (a hybrid between LCT and holograms), thermocapillary, and magneto-optics exist, but are of less relevance [9]. In addition, it is always possible to use semiconductor optical amplifiers (SOAs) as the means for switching. Table 15.3 lists relevant parameters for the most relevant switching technologies.

15.2. AMPLIFIERS

15.2.1. Erbium Doped Fiber Amplifiers

Erbium doped Fiber Amplifiers (EDFAs) are (traveling-wave, TW) laser amplifiers. An EDFA basically consists of a certain length of Er^{+} doped fiber,

Table 15.3
Characteristics of Different Switching Technologies.

	MEMS	LCT	Bubbles	Thermo-O.	Acousto-O.	Hologram
Scalability	Very high	Medium	Medium	Medium	Medium	High
Switch speed	Medium	Medium	Medium	Medium	Fast	Very fast
Reliability	Moving parts	Good	Good	Good	Good	Good
Insertion loss	4×4: 3 dB 16×16: 7 dB	2.5 to 6 dB, up to 32×32	32×32: 5 dB	8×8: 8 dB	1×2: 2.5 dB	Low
Power consumption	Medium	Very low	Medium	High	Medium	Medium, but high voltage

typically 10 to 100 m, and a pump laser diode (PLD) for providing the laser pump energy. Further components are noise-reduction filters and optical isolators for blocking backreflection.

EDFAs can amplify high numbers of wavelengths simultaneously. They provide low amplifier noise, high small-signal gain, and high-output power. Due to their analog behavior, interchannel crosstalk and the accumulation of detrimental effects (CD, PMD, nonlinearity) must be considered. In order to monitor ILAs based on EDFAs, an optical supervisory channel (OSC) is necessary. This management channel uses an outband wavelength (1510 nm according to ITU-T G.692; also 1310 nm or 1625 nm). The OSC creates an intra-WDM-system DCN (data communications network), which can be based on IP or OSI IS-IS routing.

If ultrabroadband WDM spectra have to be amplified, several EDFAs with different co-doping can be used. These are connected in parallel using appropriate band splitters (e.g., C/L-band splitters). For metro WDM ring systems, per-group amplification is also available [3], which supports flexible, reliable, and cost-efficient WDM rings with complex traffic patterns.

Similar fiber amplifiers for the wavelength range around 1310 nm have been developed, based on Praesodymium-doped fibers (PDFAs). Due to cost and performance (noise) reasons, they have not been widely adopted.

EDFAs can make use of two consecutive gain blocks. This enables low-noise design, very-high-output power, and midstage access. Midstage access can be used for connecting CD compensators; the corresponding insertion loss can be as high as 10 dB.

For 1450 to 1525 nm (so-called S-band), Raman amplifiers can be used. These amplifiers make use of the nonlinear effect of stimulated Raman scattering (SRS) inside the transmission fibers. This turns the transmission fiber into a distributed fiber amplifier. Raman amplifiers can also be used as low-noise preamplifiers in

Table 15.4
Characteristics of Optical Amplifiers.

	SLA	EDFA	PDFA	Raman	Brillouin	FWM
Material	Semi-conductor	Er^+-doped Fiber	Pr^+-, Nd^+-doped Fiber	GeO_2 in Fiber	Fiber	Fiber
Bandwidth	50–70 nm	10–40 nm	>10 nm	40 nm	0.001 nm	1 nm
Lambda range	0.8–1.6 µm	1.55 µm	1.3 µm	f(Pump)	f(Pump)	f(Pump)
Max. gain	20–40 dB	30–50 dB	30–50 dB	20–30 dB	20–30 dB	30–40 dB
Satturated gain	10 dBm	15–20 dBm	10 dBm	20 dBm	0 dBm	?
Coupling loss	5–6 dB	0–1 dB	0–1 dB	0–1 dB	0–1 dB	0–1 dB
Polarization dependence	High	Very low	Very low	Medium	Medium	High
Pump power opt./electr.(e)	(e) 0.5 W	0.1–0.2 W	0.2–0.6 W	0.5–3 W	0.01 W	30–70 W
Pump wavelength	—	800, 980, 1480 nm	1015 nm	f(Signal)	f(Signal)	f(Signal)
Active length	300 µm	10–100 m	10–100 m	0.2–100 km	10 km	1–10 km
Integrated	Yes	No	No	No	No	No
Noise	Medium	Low	Medium	Low	High	Very low

single-span configurations with very high link loss (>40 dB), or as booster and preamplifiers in multispan configurations.

Further amplifier technologies include SLAs and some other nonlinear fiber effects. SLAs are semiconductor TW amplifiers. Unlike EDFAs, they can be integrated but are lacking due to problems concerning noise and polarization dependence. Among the fiber nonlinear effects, FWM (four-wave mixing) was seriously considered. FWM leads to ultra-low-noise parametric amplifiers with noise figures as low as $F = 3$ dB [10]. However, they suffer from problems with efficient and broadband excitation.

Optical amplifiers decrease the optical signal/noise ratio (OSNR) and hence the receive-end electrical SNR through the effect of amplified spontaneous emission (ASE). Special link budget calculations are required to analyze these links [11]. Table 15.4 lists relevant properties of optical amplifier technologies.

15.3. OPTICAL TRANSPORT NETWORK—G.709

The optical transport network (OTN) standard G.709, together with other OTN standards, was developed to solve problems with SONET/SDH transparent E2E service provisioning and interworking. In SONET/SDH networks, no transparent services can be provided since the SONET/SDH client signal overhead is always terminated. Also, management problems exist with respect to different vendor and network domains.

15.3.1. Layers in OTN

G.709 defines a hierarchical transport structure. The basics are *optical channels* (OCh), which are substructured into optical channel payload unit (OPU), optical channel data unit (ODU), and optical channel transport unit (OTU). These are defined for interdomain (network, vendor) interworking only (Fig. 15.13).

ODUs are E2E transport containers similar to SONET STS or SDH Virtual Containers. So far, three hierarchical levels have been defined: ODU1 (2.50 Gb/s), ODU2 (10.04 Gb/s), and ODU3 (40.32 Gb/s).

OTUs are the corresponding P2P transport frames that are transported over wavelengths. They are equivalent to SONET/SDH Section signals (OC-n, STM-m). They have higher bit rates: OTU1 (2.67 Gb/s), OTU2 (10.71 Gb/s), and OTU3 (43.02 Gb/s).

OTN frames contain 4 × 4080 bytes, independent from the hierarchy level. Unlike SONET/SDH, OTN has no constant frame duration.

Multiplexing of several OCh leads to the optical multiplex section (OMS) and the optical transport section (OTS). The OTS is terminated between inline amplifiers, and it can be monitored through an OSC (Fig. 15.14).

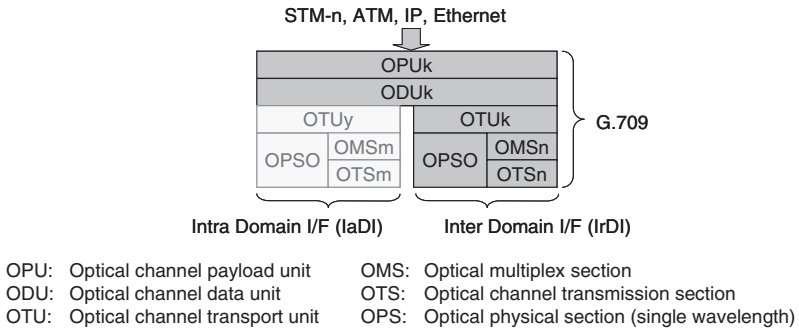


Figure 15.13 Layer in G.709.

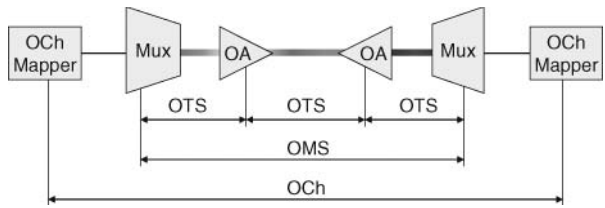


Figure 15.14 Layer above OCh (OA: Optical Amplifier).

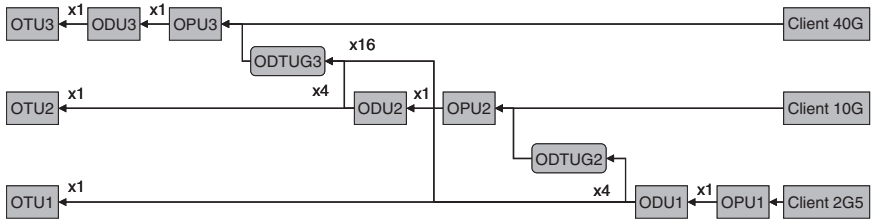


Figure 15.15 OTN Multiplexing.

Unlike SONET/SDH, OTN is not synchronized centrally. It is however possible to transport SONET/SDH synchronization (G.8251). For further details, refer to the relevant ITU standards (G.872, G.709, G.798).

A significant part of the bandwidth added between ODUs and OTUs is used by forward error correction (FEC). FEC can detect and, to a certain extent, correct bit errors. Using forward error correction (FEC), the corresponding links can support lower receive-end OSNR.

In OTN, cross-connectivity can be provided for the ODUk and OCH layers. ODUs can be cross-connected electrically. OCh cross-connect is possible electrically or optically. Performance monitoring of the OCh is only possible for electrical cross-connects. Between the ODU layers, multiplexing is defined as per Fig. 15.15.

15.3.2. Forward Error Correction

Forward error correction (FEC) adds redundant data to messages, which allows the receiver to detect and correct errors (within some bound). This leads to the acceptance of lower receive-end (O)SNR, at the cost of higher bandwidth requirements. Digital communication systems that use FEC tend to work perfectly above a certain minimum SNR and not at all below it. For ITU G.709, the FEC code used is a Reed-Solomon RS(255,239) block code. This is byte interleaved to increase burst error performance. FEC detects and corrects errors to effectively deliver a 7–8 dB improvement in SNR. The FEC check parity bytes are added when the OTUk structure is generated and are located in columns 3825 to 4080.

15.3.3. OTN Monitoring

In OTN, monitoring per layer has been defined. The OAM (operation, administration, and maintenance) functions are supported by additional path trace and loop installation. Fault location is possible with electrical and optical supervisory signals. It is possible to detect signal degradations and switch upon these

Table 15.5

Bandwidth Efficiency of SONET/SDH Transport in OTN.

Client	Bit Rate	ODU1	ODU2
10GbE WAN	9.953 Gb/s	—	~100%
SONET/SDH OC-12/STM-4	622 Mb/s	~25%	~12.5%
SONET/SDH OC-48/STM-16	2.488 Gb/s	~100%	~25%
SONET/SDH OC-192/STM-64	9.953 Gb/s	—	~100%

conditions. In ITU-T G.872, several OTN protection schemes are defined, including 1 + 1 and 1 : N path and SNC (sub-network connection) protection for the OCh and OMS layers and shared protection rings (OCh-Layer). OTN restoration is standardized in ITU-T ASON. An end-to-end monitoring scheme that was not possible using SONET/SDH is defined in OTN (tandem connection monitoring, or TCM).

15.3.4. Interworking with SONET/SDH

OTN, together with SONET/SDH, GFP, and virtual concatenation, provides high-bandwidth efficiency for data transport. In addition, Table 15.5 lists the efficiencies for transparent transport of SONET/SDH and 10GbE WAN PHY signals. 10GbE LAN PHY and 10G-FC signals can also be transported in overclocked modes.

SONET/SDH clients are mapped directly into ODUs. For client signals below OC-46/STM-16 (2.488 Gb/s), no efficient mapping is achieved. For these signals (OC-3/STM-1, OC-12/STM-4, GbE, FC), OTN can be complemented by SONET/SDH with its finer granularity (STS-1/VC-4), together with the possibility of contiguous or virtual concatenation.

Many mapping options for data and synchronous services into SONET/SDH and OTN exist, including various layers of SONET/SDH granularity. Efficient transport networks can be reduced to the STS-1/VC-4 and OC-192/STM-64/OTU2 layers (with OTU3 upcoming in 2007) [20]. The most relevant mapping options are shown in Fig. 15.16.

15.3.5. OTN Applications

OTN provides efficient high-capacity transport. As compared to SONET/SDH, new E2E services and additional functionality like common control plane for the electrical and optical layer or FEC can be provided. Besides transparent SONET/SDH transport, some OTN applications in the data and storage context can be identified.

Constant bit-rate (CBR) services have constant bit rates, for example, 2.5 Gb/s, without SONET/SDH framing. An example is InfiniBand (IB1X). Without OTN,

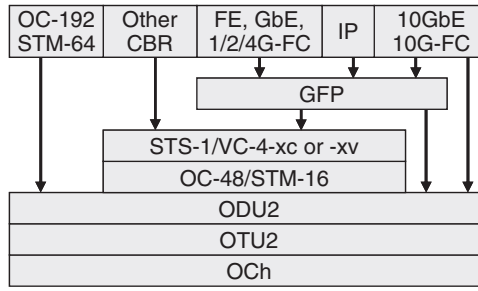


Figure 15.16 Mapping in SONET/SDH-OTN.

the only possibility for transport was dedicated WDM. OTN provides mapping of these services, including transparent transport and ODU multiplexing. Direct mapping into an OCh is also possible.

Storage area networks and distributed servers/mainframes are increasingly using transport infrastructure. In order to enable interworking and efficient transport, combinations of OTN and GFP have to be used. In addition to interworking advantages, these provide lower cost as compared to PoS.

15.4. 40G, 100G, AND HIGHER—PROBLEMS AND SOLUTIONS

15.4.1. 40G Applications

Today, there is a single relevant driver for 40G: high-performance routers. Such routers, capable of 100 Tb/s throughput, require 40G interfaces mainly because fewer but bigger pipes dramatically ease the routing between different ports and hence all functions such as link bundling, load sharing, and restoration. Fewer pipes (interfaces) are also mandatory from the viewpoint of form factor and heat dissipation: high-performance routers can no longer be built based on multiple 2G5 or 10G interfaces. The introduction of 40G is driven by technical demand rather than price.

An application area with very high bit rate demands is storage. However, no native 40G application (40G-FC, 40GbE) exists today. Hence, the only storage application that can make use of 40G is IP-storage, using FCIP, iFCP, or iSCSI.

For many large storage applications, a large number of fibers are available, leading to low-cost CWDM or even converter solutions instead of sophisticated high-end. Since no native 40G storage application exists, more or less excessive TDM is necessary in order to fill the 40G pipes. This leads to accumulated delay and jitter, which cannot be tolerated for certain storage protocols. Finally, today's

relevant storage protocols (FC, FICON) use handshaking mechanisms for flow control. For longer distances this must be complemented by costly credit buffering in order to provide high throughput. The hardware effort that is necessary for credit buffering increases linearly with increasing bit rate.

Finally, 40G is no cost-efficient technology today, mainly due to components prices and the transmission impairments. As a result, 40G is still more costly than multiple 10G.

15.4.2. Compensation of Chromatic Dispersion

Chromatic dispersion (CD) can generally be compensated using suitable techniques like dispersion compensating fibers (DCFs), lumped components like gratings, electronic dispersion compensation (EDC) at the receiver, pre-chirp at the transmitter, or nonlinear techniques like spectral inversion or soliton transmission. For examples refer to [21–23].

40G requires tight compensation of CD, including the slope. The precision of the compensation, that is, the amount of residual CD, has to be improved for 40G, as compared to 10G. This also means that dispersion compensators have to be adapted exactly to the fiber type as different types and different brands of fibers exhibit different CD, as shown in Fig. 15.17. This leads to complex sets of dispersion compensation modules and compensation schemes where fibers are compensated by means of frequently placed DCMs, with the residual dispersion being compensated by EDC.

For 40G, compensators for all relevant fiber types and brands—G.652, TrueWave-RS®, E-LEAF®, etc.—are mandatory. The relevant parameter for successful compensation is the quotient of dispersion parameter D/S and slope.

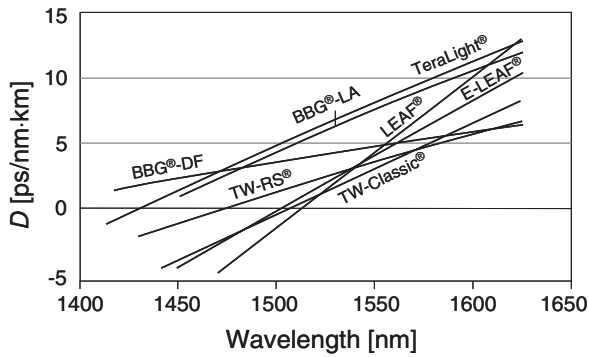


Figure 15.17 Chromatic dispersion of different ITU-T G.655 and G.656 fibers.

Table 15.6
Dispersion Characteristics of Transmission Fibers and DCF.

	D [ps/(nm·km)]	S [ps/(nm^2·km)]	D/S [nm]
G.652	16.7	0.056	298
Hi-slope DCF	–105	–0.35	300
Low-slope G.655	6.6	0.045	147
Ultra-slope DCF	–115	–0.78	147

Table 15.6 lists two pairs of fibers (standard G.652 and reduced-slope G.655) together with well-adapted compensation fibers.

For various reasons (providing the necessary power budget for the dispersion compensators, suppression of SPM and other nonlinear effects), DCMs should be placed frequently along the transmission line. Remaining residual CD can then be compensated electronically. For many metro and regional applications, network operators do not have the choice as to exactly where to place compensators. Typically, for these applications no equidistant amplifier spacing exists. Hence, the modulation scheme must be robust enough to cope with nonequidistant amplification and compensation, and with variable amounts of residual CD and uncompensated polarization-mode dispersion (PMD). In ULH, dispersion management can be used (i.e., links that consist of successive spans of fibers with positive and negative D parameter, thus providing net CD, which is close to zero; see, for example, [15]).

15.4.3. Fiber PMD

For 40G and above, PMD must be considered very closely. Depending on the fiber quality—G.652A/C and G.655A/B vs. G.652B/D, G.655C, and G.656—and the link length, PMD must even be compensated. Only for shorter distances—typically below 100 km—and on fibers with proper PMD, it can be considered by means of a simple PMD penalty if bit rates are 40G and more. PMD is described by means of a parameter D_P. This parameter has the dimension [ps/√km]. The standards ITU-T G.652.A/C, G.653 and G.655A/B define a maximum of 0.5 ps/√km; the newer standards G.652B/D, G.655C and G.656 only allow for 0.2 ps/√km. Depending on the fiber characteristics, PMD can lead to severe limitations of the maximum transmission distance. In order to avoid complex compensation, the PMD allowance of transmission systems has to be as big as possible. Typical requirements of network operators range around 10 ps. PMD compensation techniques are discussed, for example, in [24]. Due to its stochastic nature, PMD is more difficult to compensate than CD. It has been shown in several extensive audits that over 20% of installed fibers, even those installed after the year 2000, can produce PMD in excess of 1 ps/√km. [25, 26].

15.5. ROADMS—TECHNOLOGY OVERVIEW AND APPLICATIONS

Typical metro networks consist of a core layer and an access layer. The core in most cases consists of static WDM technology and SONET/SDH rings.

The access layer consists of a mixture of different technologies, in particular SONET/SDH rings (complemented by GFP), CWDM access rings and Linear Add/Drop links, and packet-oriented, native Ethernet access systems. Further access technologies like wireless and PON also exist. They have in common that their usage strongly depends on the access fiber infrastructure, the infrastructure that is used for aggregation (transparent WDM vs. SONET/SDH vs. packet-oriented), and the management requirements of the operator (SONET/SDH-like vs. IP-based management).

A cost-effective photonic core network is not realized without optical bypass at intermediate nodes. Optical bypass is provided with OADMs. With the increase of IP traffic and the requirements with respect to maximum node numbers in core rings, the optical bypass has to be reconfigurable, with single-channel add/drop capabilities. This requires ROADMs for core rings. O/E/O switching and aggregation will then move to service termination and grooming points at the edge of the access network (between access and last mile) and into some core PoPs, as a second aggregation layer.

15.5.1. ROADM Technology

A ROADM is an optical switching device, but the name is commonly used for dynamic optical layer. ROADMs allow the network management system (NMS) or a control plane to control whether wavelengths are routed through the node or to a local port where they are terminated on a transponder or client interface. ROADMs of Degree 2 can switch individual wavelengths between two ports (one aggregate port—east or west—and the local port). Aggregate ports support WDM signals; the local ports can be wavelength agnostic or specific and may support single or multiple wavelengths. This general ROADM functionality is shown in Fig. 15.18.

The principal ROADM functionality can be achieved with different switching technologies and different resulting design concepts. The corresponding ROADMs can be based on:

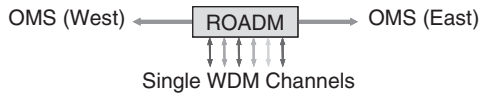


Figure 15.18 Degree-2 ROADMs allow reconfigurable single-channel add/drop.

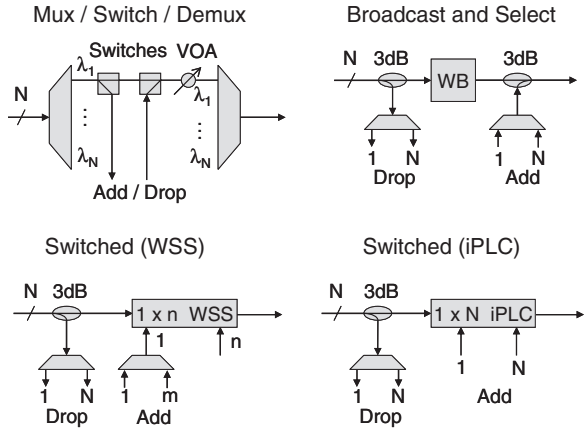


Figure 15.19 ROADM architectures overview.

- Discrete switches or switch matrix plus filters (Mux/Switch/Demux)
- Wavelength blockers (WB)
- Integrated planar lightwave circuits (iPLC)
- Wavelength-selective switches (WSS)

Figure 15.19 shows the principal ROADM technologies. With the exception of mux-switch-demux design, the devices are typically implemented in broadcast-and-select optical architectures with passive splitters in the pass-through path.

A relevant attribute of ROADM technology is the integration of multiplexing/demultiplexing and switching into a single component. This integration can significantly lower pass-through losses when compared with multiple discrete components. Lower loss results in improved OSNR and larger ring sizes with more nodes.

iPLC technology offers integration of AWG multiplexers/demultiplexers, switches, taps, monitoring diodes, and VOAs (variable optical attenuators). This offers the highest degree of integration today. The advantages and disadvantages of the ROADM architectures are listed in Table 15.7.

ROADMs based on iPLCs offer several advantages over the other technologies, notably, comparatively low insertion loss (~7.5 dB express path plus splitter), high integration, and the possibility to be upgraded to multidegree connectivity. The monitoring and VOA capabilities allow power leveling, which is the basis for extended scalability of the number of nodes without significant Q-penalty.

WSSs enable banded or uncolored add/drops. As the demultiplexer is not included on the ROADM module, modular upgrade of the multiplexer/demultipexer structure is possible. Furthermore, a WSS-based Degree-4 ROADM can be built

Table 15.7
Feature Comparison of ROADM Architectures.

ROADM architecture option Criterion	M-S-D	WB	iPLC	WSS
Insertion loss	—	+	+	o
Number of drop ports	40	40	40	5–9
Integration (VOAs, taps)	—	VOAs	+	+
Higher-degree nodes with λ-patching	no	yes	yes	n/a
Higher-degree nodes without λ-patching	no	no	no	yes
Colored, λ-specific adds/drops	yes	yes	yes	n/a
Uncolored drops	no	no	no	yes
Banded drops	no	no	no	yes

when 1 : 4 splitters are used. Degree-4 nodes can be used to interconnect two rings or to build meshed networks. Finally, integrated power monitoring can be used for measurements of the OSNR. WSSs are based on LCT or MEMS, as compared to AWG-based iPLCs.

Ways to lower ROADM cost include *modularity*. The initial service requirement for a network is typically much less than its maximum capacity. Modularity allows service providers to delay common equipment CapEx until they are needed, thus aligning CapEx associated with the delivery of a service more closely with the revenue generation. In fact, many networks never reach their maximum capacity. In the absence of modularity, much of the initial investment would then effectively be lost.

In addition to the pay-as-you-grow advantage, modularity also considers the fact that in most networks there is a certain amount of static traffic load. This is even true for IP networks, where certain main routes are established as permanent MPLS circuits. The capacity gain (or the corresponding cost decrease per bit/s) of a network which can be achieved by providing an increasing amount of flexibility is clearly limited, as is shown in Fig. 15.20. The diagram shows the increase in total network capacity for meshed and ring networks (i.e., for different degrees of meshing) over the relative amount of flexibility provided. The first 25 to 50% of added flexibility lead to a significant increase of total network capacity in the range 30 to 40%, whereas further increase toward full flexibility contributes no longer to significant capacity gain.

Advanced ROADMs support nodes with connectivity higher than Degree-2, which requires the ROADM to have more than one pass-through port. These higher-degree ROADMs can be used to physically connect meshed networks. The corresponding higher-degree ROADMs are sometimes also referred to as (all-) optical cross connects, or OXCs (i.e., without O/E/O conversion).

To provide SONET/SDH-like wavelengths cross-connect functionality, multi-degree ROADMs need to be upgradeable in-service in a modular way. All tech-

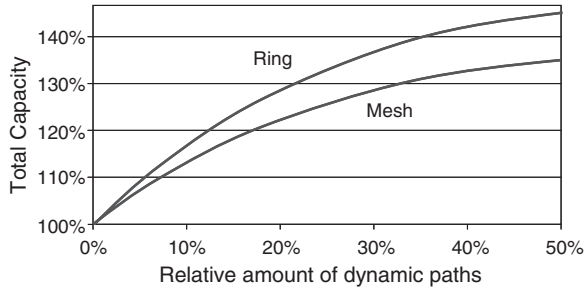


Figure 15.20 Advantage of providing reconfigurability in meshed networks.

nologies listed in Table 15.7 support higher-degree ROADMs, but differences exist with respect to the necessity of wavelength-selective patching, which is a function of the underlying switching technologies.

A critical parameter of multidegree ROADMs is wavelength blocking. In a ring-interconnecting ROADM, wavelength blocking occurs if wavelengths (or groups of wavelengths) of one ring are cross-connected into the other ring where the same wavelengths are already in use for other connections. Wavelength blocking can be a severe problem for meshed networks, and it almost always prevents the possibility to transparently connect several access rings to a core ring. Wavelength blocking hence leads to the necessity of 3R regenerators or transparent, all-optical wavelength converters in a network.

In the access, where access rings with hubbed traffic join core rings in a reconfigurable add/drop site, 3R regenerators will be used. Access rings may be based on CWDM rather than DWDM, so that technology conversion between both is necessary. Even if DWDM access rings are used, these most likely will make use of low-cost, low-performance interfaces so that 3R regeneration again is necessary. In addition, regeneration provides demarcation between different network domains. A nonblocking degree-4 ROADM, which flexibly connects a core ring to an access ring, can be designed by combining two Degree-2 ROADMs with wavelength converters (3R regenerators) in between them; see Fig. 15.21.

The ROADM shown in Fig. 15.21 is useful for connecting access rings with hubbed traffic to core rings. In addition, there is the desire to provide *transparent* wavelength conversion within core or regional networks. These transparent conversions help omitting expensive 3R regenerators. In addition, certain all-optical techniques allow truly transparent conversions, thus being ready for 10G, 40G, or any other bit rate in advance.

Several techniques were investigated for transparent wavelength conversion, which all make use of nonlinearity as the mechanism to generate new frequencies.

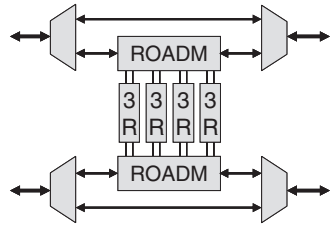


Figure 15.21 Ring interconnection using two Degree-2 ROADMs.

Suitable nonlinear effects are parametric mixing (four-wave mixing) within dispersion-shifted fibers, a similar effect in quadratic nonlinear (bulk) media like $LiNbO_3$, or frequency shifting in semiconductor optical amplifiers (SOA) that are driven into the nonlinear regime.

The use of parametric mixing for wavelength conversion (also referred to as spectral inversion) has been well known for more than two decades [10, 17]. Since it exhibits several problems, it never matured to the level necessary for commercial products. The use of SOAs was more recently described as an interesting alternative [40].

ROADMs offer significant advantages over static solutions for a number of applications. These include any-to-any connectivity and single-channel add/drop in large rings (which significantly reduce OpEx), dynamic network planning, reconfiguration, and restoration. Centralized Layer-3 networks are also enabled by this technology.

REFERENCES

1. Saleh, A. A. M., and J. M. Simmons. 1999, December. Architectural principles of optical regional and metropolitan access networks. *IEEE Journal of Lightwave Technology* 17, no. 12: 2431–2448.
2. Grobe, K. 2004, May/June and September/October. Optical metro networking. Telekommunikation Aktuell, 58. Jahrgang, No. 5/6 and 9/10.
3. ADVA AG Optical Networking. *FSP 3000 Introduction*. http://www.advaoptical.de/adva_products.asp?id=133.
4. Barry, M., et al. 2000, May. A classification model of network survivability mechanisms. Proceeding 5. ITG-Fachtagung Photonische Netze, Leipzig, May 2004.
5. Arijs, P., et al. Architecture and design of optical channel protected ring networks. *Journal of Lightwave Technology* 19, no. 1:11–22.
6. Gerstel, O., and R. Ramaswami. 2000, March. Optical layer survivability: A services perspective. *IEEE Communications Magazine* 38, no. 3:104–113.
7. Fang, X., R. Iraschko, and R. Sharma. 1999, August. All-optical four-fiber bidirectional line-switched ring. *IEEE Journal of Lightwave Technology* 17, no. 8:1302–1308.
8. European Patent EP 1371163B1. 2002, March. Selbstheilende Ringstruktur zur Optischen Nachrichtenübertragung im Wellenlängenmultiplex und Add/Drop-Multiplexer hierfür.

9. Optical switching devices. 1999, December. Report ON-2 San Antonio: Strategies Unlimited.
10. Löcherer K.-H., and C.-D. Brandt. 1982. *Parametric electronics*. Berlin: Springer.
11. Haus, H. A. 1998, November. The noise figure of optical amplifiers. *IEEE Phot. Tech. Let.* 10, no. 11:1602.
12. Gnauck, A. H., and P. J. Winzer. 2005, January. Optical phase-shift keyed transmission. *IEEE Journal of Lightwave Technology* LT-23, no. 1:115–130.
13. Agrawal, G. P. 1992. *Fiber-optic communication systems*. New York: John Wiley & Sons.
14. Agrawal, G. P., and Z. M. Liao. 2001, July. Role of distribited amplification in designing high-capacity soliton systems. *Optics Express* 9, no. 2:66–71.
15. Hasegawa, A. 2002, February. Optical solitons in fibers for communication systems. *Optics & Photonics News*, OSA.
16. Mollenauer, L. F., R. H. Stolen, and J. P. Gordon. 1980. Experimental Observation of Picosecond Pulse Narrowing and Solitons in Optical Fibers. *Physical Review Lett.* 45, no. 13:1095.
17. Agrawal, G. P. 1995. *Nonlinear fiber optics*. 2nd ed. San Diego: Academic Press.
18 Generic framing procedure and data over SONET/SDH and OTN. 2000, May. *IEEE Comm. Magazine* 40, no. 5, issue on GFP.
19. Bonenfant, P., and A. Rodriguez-Moral. 2002, May. Generic framing procedure: The catalyst for efficient data over transport. *IEEE Comm. Magazine* 40, no. 5:72.
20. Eilenberger, G., et al. 2004, March. *OTN—Technical trends and assessment.* 5. ITG Fachtagung Photonische Netze, Leipzig, und WDM Conference, Cannes.
21. Ohm, M., T. Pfau, and J. Speidel. 2004, May. Dispersion compensation and dispersion tolerance of Optical 40 Gbit/s DBPSK, DQPSK, and 8-DPSK Transmission Systems with RZ and NRZ Impulse Shaping. Proc. 5. ITG-Fachtagung Photonische Netze, Leipzig.
22. Royset, A., et al. 1996. Linear and nonlinear dispersion compensation of short pulses using mid-span spectral inversion. *IEEE Phot. Tech. Lett.* 8, no. 3:449.
23. Taga, H., et al. 1994. Performance evaluation of the different types of fiber-chromatic-dispersion equalization for IM-DD ultralong-distance optical transmission. *IEEE Journal of Lightwave Technology* LT-12, no. 9:1616.
24. Buchali, F., and H. Bülow. 2004, April. Adaptive PMD compensation by electrical and optical techniques. *IEEE Journal of Lightwave Technology* LT-22, no. 4.
25. Peters, J., et al. 1997, September. Bellcorés fiber measurement audit of existing cable plant for use with high bandwidth systems. NFOEC San Diego, Calif.
26. Barcelos, S., et al. 2005, March. Polarization mode dispersion (PMD) field measurements—*An audit of brazilian newly installed fiber networks*. OFC2005, Anaheim, Calif.
27. Shankar, H. 2004. *Duobinary modulation for optical systems*. White Paper, Inphi Corporation.
28. Bhandare, S., D. Sandel, B. Milivojevic, A. F. A. Ismael, A. Hidayat, and R. Noe. 2004, May. *2 × 40 Gbit/s RZ DQPSK Transmission.* Proceeding 5. ITG-Fachtagung Photonische Netze, Leipzig.
29. Serbay, M., C. Wree, and W. Rosenkranz. 2004, May. *Kostengünstige Realisierung einer robusten Übertragung mit dem DQPSK-Modulationsformat einschließlich Vorkodierung.* Proceeding 5. ITG-Fachtagung Photonische Netze, Leipzig.
30. Zhu, Y., et al. 2004, March. 1.6 bit/s/Hz orthogonally polarized CSRZ-DQPSK transmission of 8 × 40 Gbit/s over 320 km NDSF. OFC2004, Anahein, Calif.
31. Nakamura, S., et al. 2000, April. Demultiplexing of 168-Gb/s data pulses with a hybrid-integrated symmetric Mach-Zehnder all optical switch. *IEEE Photonics Tech. Lett.* 12:425–427.
32. Schubert, C., et al. 2001, November. 160-Gb/s Polarization Insensitive All-Optical Demultiplexing Using a Gain-Transparent Ultrafast Nonlinear Interferometer (GT-UNI). *IEEE Phot. Tech. Lett.* 13, no. 11:1200–1202.
33. Sokoloff, J. P., and P. R. Prucnal. 1993, July. A terahertz optical asymmetrical demultiplexer (TOAD). *IEEE Phot. Tech. Lett.* 5, no. 7:787–790.

34. Gordon, J. P., and L. F. Mollenauer. 1990. Phase noise in photonic communications systems using linear amplifiers. *Optics Letters* 15, no. 23:1351.
35. Kim, H., et al. 2003, February. Experimental investigation of the performance limitation of DPSK systems due to nonlinear phase noise. *Photonics Tech. Letters* 15:320–22.
36 ITG-Fachausschuss 5.3 Optische Nachrichtentechnik: Breitbandige Kommunikationsnetze mit hoher Qualität: flexibel effizient intelligent. 2005, VDE/ITG-Positionspapier service@vdecom.
37. Raybon, G., et al. 2006, March. 10 × 107-Gbit/s Electronically multiplexed and optically equalized NRZ transmission over 400 km, OFC 2006, Anaheim, Calif., PDP 32.
38. Derksen, R. H., et al. 2006, July. 100 Gbit/s Ethernet for true end-to-end carrier-grade Ethernet networks. NOC Berlin.
39. Antonaides, N., et al. 2006, July. An architecture for a wavelength-interchanging cross-connect utilizing parametric wavelength converters. *IEEE Journal of Lightwave Technology* 17, no. 7:1113–1125.
40. Leuthold, J., et al. 2003, November. Non-blocking all-optical cross connect based on regenerative all-optical wavelength converter in a transparent demonstration over 42 nodes and 16,800 km. *IEEE Journal of Lightwave Technology* 21, no. 11:2863–2869.
41. U.S. Communications Infrastructure at a Crossroads. Goldmann Sachs McKinsey & Company. August 2001.
42. St. Arnaud, et al. 2003, January. Customer Controlled and Managed Optical Networks. http://www.canarie.ca/canet4/library/c4design/customer_controlled.pdf.
43. Oki, E., et al. 2002. A heuristic multi-layer optimum topology design scheme based on traffic measurement for IP + photonic networks. OFC, paper TuP5.
44. Pongpaibool, P., et al. 2002. *Handling IP traffic surges via optical layer reconfiguration*. OFC, paper ThG2.
45. Wei, J. 2002. *IP over WDM network traffic engineering approaches*. OFC, paper TuP4.
46. Foster, I., and C. Kesselmann. 1998. *The grid—Blueprint for a new computing infrastructure*. San Francisco: Morgan Kaufmann Publishers.
47. What is DRAGON? http://dragon.maxgigapop.net.
48. The RAY product portfolio. http://www.movaz.com/Products.aspx.

Case Study National LambdaRail Project

Provided in part by Cisco Systems

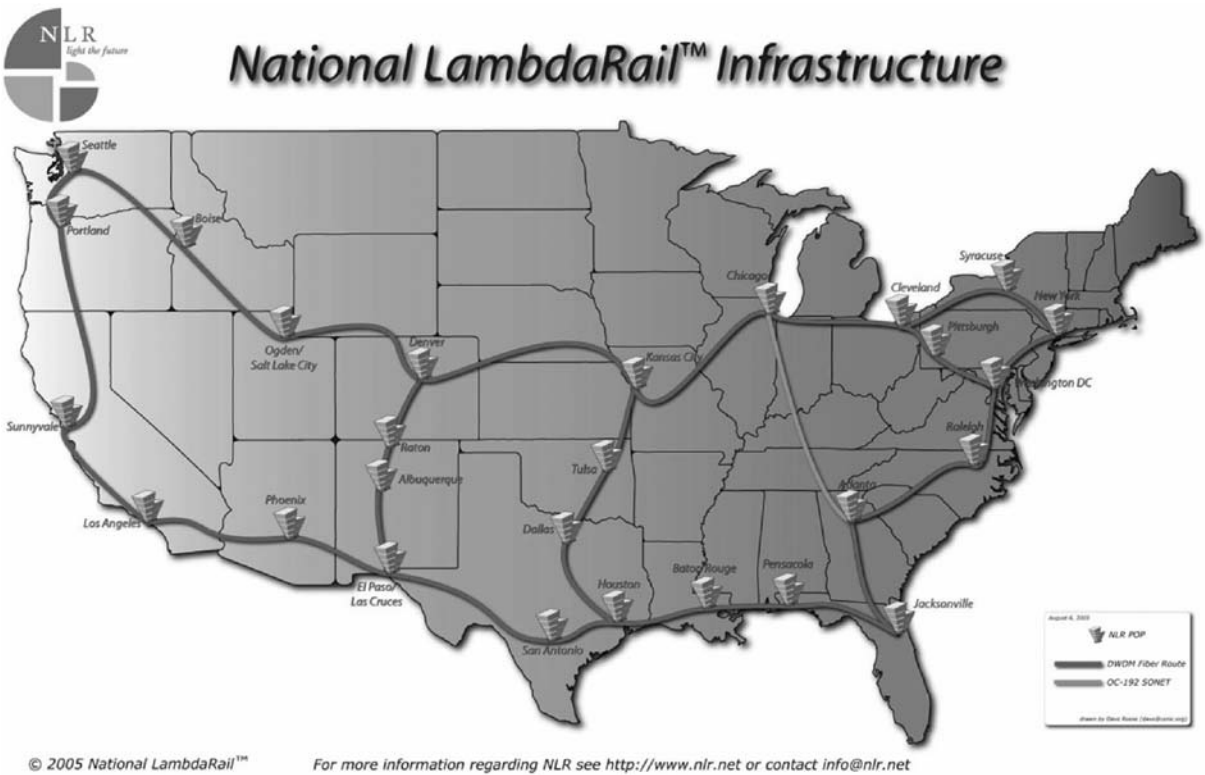
One of the widely cited examples of the powerful capabilities of optical networking is the National LambdaRail project (http://www.nlr.net/), a high-speed fiber-optic computer network in the United States owned and operated by a consortium of universities. Its primary goal is the support of terascale computing grids, though it is also used as a testbed for experimentation with next-generation large-scale networks. The National LambdaRail network was developed to help bridge the gap between leading-edge optical network research (beyond the bounds of current Internet backbones) and computationally or bandwidth-intensive application research projects (one recent example is the FCC Rural Health Care Pilot program (http://www.fcc.gov/cgb/rural/rhcp.html), intended to provide telemedicine services under the National Healthcard Delivery Initiative). As illustrated in the figure, National LambdaRail's intracity backbone consists of DWDM equipment presently carrying up to 10 Gbit/s per wavelength. This network must balance the conflicting objectives of offering a highly flexible, leading-edge design to promote creativity among its end users, while at the same time ensuring a highly stable, reliable connection for collaborating research institutions.

Various regional subnetworks comprise National LambdaRail, for example, the Florida LambdaRail (FLR). Operational since March 2005, this network interconnects 10 major research institutions and spans the entire geography of the state of Florida. The network design employs the Cisco 15454 Multiservice Transport Platform (MSTP) DWDM backbone, as well as many Cisco Catalyst 3750 switches and 7600 series routers. Optical amplifiers provide extended distance capability as required, and ROADMs provide the capability to automate many network maintenance functions, topology discovery, and the addition of new wavelengths. The 15454 platform supports both SONET and Ethernet traffic,

which accommodates both production and research traffic running side by side in the same network. All network operations for a given institution can be combined into a single fiber path, which is then subdivided and segmented into VPNs as required. Optical power levels are monitored on a per-channel basis, making it possible to implement power balancing and take the best advantage of optical amplifier placement. By deploying a shared IP infrastructure, participants have been able to document cost savings compared with OC-12 services. Each member institution has a 10 Gbit/s connection to the network and a 1 Gbit/s backup link; participants share a 10 Gbit/s access point in the Atlanta node of National LambdaRail.

REFERENCE

"Florida LambdaRail powers advanced academic research and communication," Cisco Systems case study, 2006, http://www.cisco.com/en/US/products/hw/optical/index.html.


National LambdaRail™ Infrastructure
NLR
light the future
Seattle
Portland
Boise
Ogden/
Salt Lake City
Denver
Kansas City
Chicago
Cleveland
Syracuse
Pittsburgh
Sunnyvale
Raton
Albuquerque
Tulsa
Phoenix
Los Angeles
Dallas
El Paso/
Las Cruces
Houston
Baton Rouge
Pensacola
Jacksonville
San Antonio
Raleigh
NLR POP
DWDM Fiber Route
OC-192 SONET
© 2005 National LambdaRail™
For more information regarding NLR see http://www.nlr.net or contact info@nlr.net

This page intentionally left blank

Case Study Optical Networks for Grid Computing

Courtesy of Nortel Optical Networks

Application: Connect two principal research centers 80 km apart at ultra-broadband data rates (10 Gbit/s) to enable a grid computing initiative.

Description: As many scientific research labs expand on a global scale, collaborative resource sharing between different parts of the same organization, or between different research institutions, has become a key enabler for new scientific discoveries. The concept of interconnecting computer systems over an infrastructure reminiscent of the electrical power utility grid has become a reality in recent years through the availability of high-bandwidth fiber-optic connectivity. A large public sector national research laboratory in Europe was faced with the challenge of interconnecting many of their facilities so that they could share resources (storage, processing power, network bandwidth) and function in essence as a large, distributed supercomputer. These facilities included a major nuclear physics facility, the national space agency and astronomy labs, several universities, and research hospitals, all of which generated on the order of several terabytes of new information each year. The first step in building this grid was to enhance the legacy network connections between two of their principal locations, which were limited to 2 Mbit/s, while their objective was for both locations to behave as if they were part of the same local area network.

The grid backbone was upgraded with a 32-wavelength dense WDM solution (the Nortel Optera Metro 5200), in order to allow several virtual networks to operate over the same fiber using different wavelengths. To provide the necessary reliability, protection switching was enabled on a per-wavelength basis. The inherent security of the optical network and the ability to detect when fibers were connected or disconnected from the equipment provided an additional layer

of security for this application. Additional multinode firewall functions were incorporated into a central routing switch (Nortel ERS 8600) in the core of the resulting 10 Gbit/s Ethernet LAN. The resulting network allowed researchers to build a computing environment with a performance (in floating-point operations per second) on a par with the top 500 supercomputers in the world.

16

Passive Optical Networks

Klaus Grobe
ADVA Optical Networking, 82152 Martinsried, Germany

16.1. PASSIVE OPTICAL NETWORKS

16.1.1. Introduction

Passive optical networks (PONs) are an access technology linking the central office (CO) or headend of a service provider to the customer premises (CP) or cabinet for fiber to the home (FTTH), business (FTTB), or curb (FTTC) applications. PONs are generally not used for metro core or long-haul applications.

PON technology allows the service provider to share the fiber cost of running fiber from the CO to the premises among many users—usually up to 32 locations. As shown in Fig. 16.1, the fiber runs from the CO to a centralized distribution point; then fiber laterals extend from this point to each customer location. The extension of the fiber is done via passive optical splitters/couplers or filters at the distribution point.

PONs do not require any power requirements in the outside plant to power the filters or splitters, thereby lowering the overall operational cost and complexity. Because the single fiber is shared in a tree (or ring) technology, the high-cost capital deployment of fiber is lower for several kilometers than if the carrier were to deploy individual fibers to each location.

PONs are the basis for broadband access networks, enabling high-speed Internet access, digital TV broadcast (IPTV), video on demand (VOD), and others. As compared to copper-based technologies like xDSL, higher bandwidths (up to several Gb/s) and higher distances (up to 10 s of km) are possible. In addition, PONs do not suffer from electromagnetic interference (EMI), which is the case in xDSL on twisted-pair cables due to crosstalk between different users. Figure 16.2 gives an overview on the bandwidths of several access technologies.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

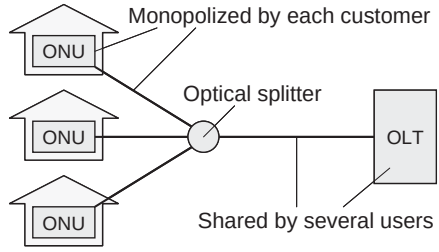


Figure 16.1 Principle of FTTH/PON access (OLT: optical line termination, ONU: optical network unit).

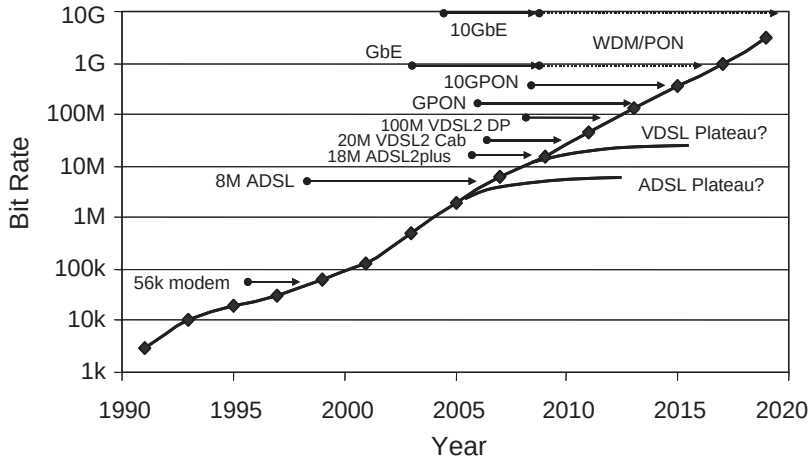


Figure 16.2 Comparison of access technologies with respect to bandwidth.

Combinations of (10)GbE and WDM/PON can exceed copper-based solutions by several orders of magnitude.

The first PON field trials were conducted in 1991 in Europe [1]. Since then, PON deployment has differed significantly in the three main regions of North America, Europe, and Asia-Pacific. In North America and Asia-Pacific, PON deployment clearly exceeded deployments of active optical networks (AONs), that is, active, dedicated point-to-point (P2P) systems. In Europe, the situation was vice versa in 2005, as can be seen from Fig. 16.3.

AON deployments stem from the fact that dedicated P2P can offer higher capacity than shared access and can thus provide better future-proofing. In addition, dedicated access offers higher security, upgrades only affect one customer,

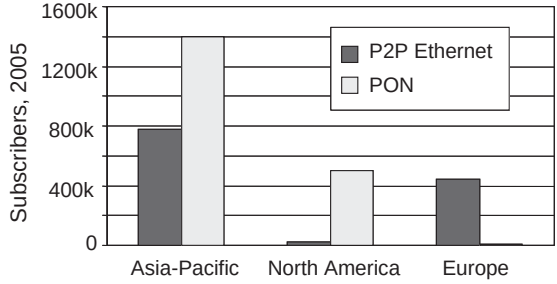


Figure 16.3 Comparison of P2P Ethernet and PON access (2005).

distances are significantly increased (up to 120 km), and service and network management can be improved.

Around 2005, PON deployment was most advanced in Asia-Pacific, with the Japanese incumbent network operator NTT the international market leader. North America was in second place, and there were only comparatively few PON deployments in Europe (but massive AON deployment instead).

16.1.2. PONs and Optical Access Networks

PONs are the basis of optical access networks (OANs) as defined in ITU-T G.902, and of hybrid access networks (hybrid fiber coaxial—HFC—networks). In G.902, the OAN as used for xDSL is split into an optical distribution part that is terminated in an optical nework unit (ONU), and a customer-facing access part using copper-based twisted pairs (unshielded/shielded twisted pair, UTP, STP).

PONs are typically configured in trees where the legs form optical distribution networks (ODNs). The different degrees of optical vs. electrical access are summarized in Fig. 16.4 for the different FTTX access scenarios. According to Fig. 16.4, the reference points for the optical access network are the service-network interface (SNI) and the user-network interface (UNI). They are described in the PON standards (ITU-T G.983, G.984).

16.1.3. Basic PON Structure and Function

In PONs, several customers are connected to a central office (CO) via a passive fiber-optic infrastructure. This infrastructure splits into single-mode fibers and a passive splitter (coupler in the reverse direction); the splitting ratios are 1 : 16 to 1 : 64. PONs work bidirectionally with data rates sufficient for broadband triple-play services. Downstream transmission usually uses time domain multiplex (TDM) across the customers, and upstream makes use of one of several multiple-access technologies.

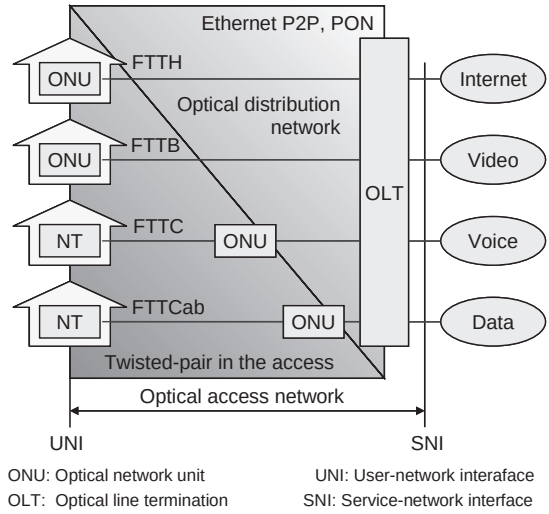


Figure 16.4 FTTX access in an OAN. FTTH/FTTB/FTTC/FTTCab: Fiber to the Home/Building/Curb/Cabinet.

Main PON components are:

- OLT: optical line termination (in CO, service provider headend)
- Optical splitter (passive, e.g., in cabinets)
- ONU: optical network unit (at CP, or in cabinet etc.), also referred to as ONT (Optical Network Termination)
- ODN, comprises of single-mode fibers and passive splitters

The OLT is the interface between the access network and the backbone. It is responsible for the optical transmission and reception into/from the PON. From the OLT, the PON extends via the passive splitter to the customer locations. The OLT is also responsible for the enforcement of the MAC protocol for upstream bandwidth arbitration and the coupling of the distribution network with the ATM transport network (if an ATM backbone is used). Optionally, switching or cross-connection is also provided to relieve the transport network of switching responsibilities.

The passive optical splitter is located in a socket or (outdoor) cabinet. Splitting ratios of 1 : 16 up to 1 : 64 can be realized, depending on the PON technology used. Lower ratios like 1 : 4 are also possible, if required. From the splitter, the PON extends to the CPs where it is terminated in ONUs.

In the ONU (ONT), the optical signals are converted back into electrical signals of the corresponding formats (e.g., POTS, 10/100 bT). Like the OLT, the ONU is an active component that requires power supply. The ONU cooperates

with the OLT in order to control the power transmitted from the residence to the carrier facility. It is also responsible, in cooperation with the OLT, for the enforcement of the MAC protocol for upstream bandwidth arbitration. The ONU acts as the residential gateway, coupling the distribution network with the in-home network medium.

The ODN (optical distribution network) is comprised of single-mode optical fibers and the passive optical components (optical splitters). It offers one or more optical paths between one OLT and one or more ONUs. The difference in attenuation (path loss) between the OLT and any two customer locations must usually be limited to 15 dB.

16.1.4. Upstream PON Access

A basic question in PONs is how the access in the *Upstream* (US), *from* the ONUs *to* the OLT, is to be organized. The problem is that N customers (with $N = 16 \ldots 64$) need to share one fiber (between the splitter and the OLT), and potentially a single wavelength as well.

In PONs, the usual technologies for multiple access (to a common recource) can be used, that is, TDMA, WDMA, SCMA, and CDMA. Today, however, only TDMA and WDMA are relevant to certain extents. The *Downstream* (DS) almost never is a problem since simple TDM (or WDM) schemes can be used.

In TDMA (time domain multiple access), the upstream is shared by the N customers through allocation of dedicated time slots per customer. This has to be done by the OLT and the ONUs, providing fairness between the customers and avoiding, where possible, collisions. In ATM PONs, *ranging* is used to measure the distance between OLT and each ONU in order to avoid collisions of upstream cells.

Advantages of TDMA are the use of identical optical sources at the CPs, and the requirement for only a single photo detector (photodiode, PD) at the CO. Disadvantages include the need for high-speed light sources (CP) and receivers (CO), which all have to operate at the aggregate bit rate. Also, a ranging protocol may be required, which can lead to an additional access delay. The TDMA principle is schematically shown in Fig. 16.5.

Figure 16.5 TDMA upstream.

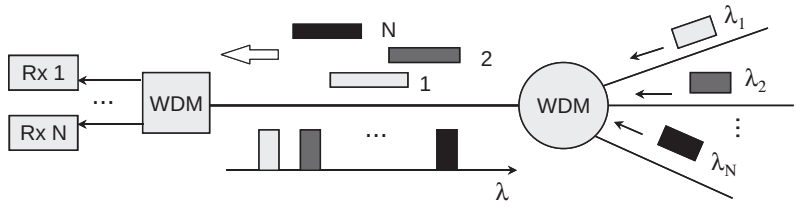


Figure 16.6 WDMA upstream.

WDMA (wavelength domain multiple access) uses different wavelengths for the upstream channels (and also different wavelengths for the downstream). Since all customers have a dedicated wavelength, no collisions can occur, and highest security is provided through physical separation. Also, customer wavelengths are independent from other customers, thus providing a certain degree of transparency and easier (capacity) upgrades. All components (especially in the CO) only need to run at the customer bit rate, not at the aggregate bit rate.

Disadvantages of WDMA include the need for dense WDM (DWDM), including the problems with temperature instability and the need for wavelength stabilization and costly DFB lasers (distributed feedback). There is no components sharing in the CO (WDMA requires *N* PDs, etc.), so that only the fiber between OLT and splitter is shared. WDMA is shown in Fig. 16.6.

A version of WDMA is *spectrum slicing*, which avoids some of the drawbacks discussed above. With spectrum slicing, the ONUs use ultra-broadband light-emitting diodes (LEDs) instead of DFB DWDM lasers. WDM demultiplexing in the downstream and WDMA in the upstream are provided through high-definition (HD) passive optics in the passive node and the CO. This way, the advantages of WDMA can be maintained, and identical, low-cost optical sources can be used in the CPs. In addition, no wavelength stabilization of active components is necessary.

On the other hand, HD WDM components are necessary (temperature drift?), and the use of LEDs leads to poor power budgets as compared to the laser-based WDMA approach as discussed above. Also, spectrum slicing does not enable components sharing in the CO. This WDMA approach is shown in Fig. 16.7.

SCMA (sub-carrier multiple access) is an alternative to TDMA/WDMA. Subcarrier multiplexing (SCM) was studied extensively in the late 1980s as a means for (analog) video transmission in the CATV context. SCMA is based on using dedicated RF subcarriers (in the several GHz range) for each customer and to transmit these on a common, high-speed, analog laser wavelength. Since different subcarriers are used, the upstream wavelengths can be combined in a simple passive splitter. In the CO, the subcarriers are demultiplexed behind the PD with a μWave splitter. SCMA is shown in Fig. 16.8.

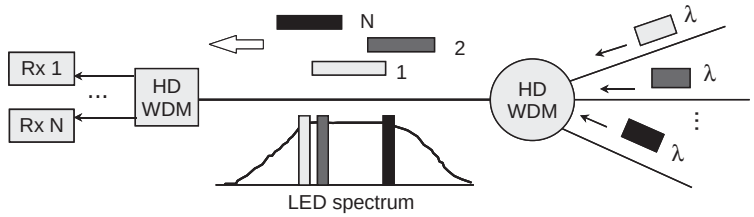


Figure 16.7 WDMA upstream—spectrum slicing.

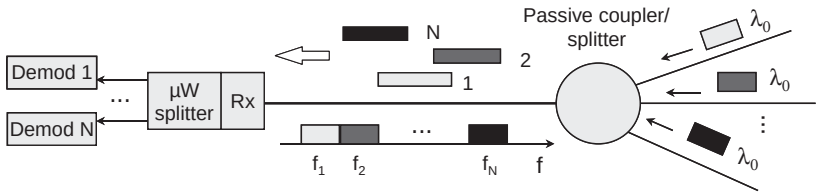


Figure 16.8 SCMA upstream.

Advantages of SCMA include the use of identical optical sources in the CPs and the need for only one PD in the CO, together with low-speed baseband transmitters and receivers. Like WDMA, SCMA provides independence of the per-customer channels. Disadvantages include the need for high-speed (analog) light sources and photodiodes, and the RF modulators and demodulators.

The latest alternative for multiple PON upstream access is CDMA (code division multiple access). CDMA is now in massive use in mobile telephony—UMTS—where it is also used to overcome some of the problems of the wireless transmission channel.

In CDMA, each (customer-specific) channel is multiplied with a high-speed coding sequence (e.g., M-sequences, Gold-sequences). The Baud rate of these coding sequences can be in the range of $100/T$, with T the bit rate of one channel. The time slots of the high-speed sequence are referred to as chips; the Baud rate is then called chip rate. For multiple access, the sequences need to be orthogonal; that is, the product of any two signals (channels) multiplied with two different sequences is always zero within one time slot T. Then, the different channels can be demultiplexed in the CO by splitting the received (multiplexed) signal and again multiplying all subsignals with the respective code sequence. Like WDMA/SCMA, CDMA does not require a complicated MAC protocol. CDMA has benefits for relatively high numbers of low bit rate connections. On the other hand, it is difficult to find optical orthogonal codes (OOCs). CDMA is shown in Fig. 16.9.

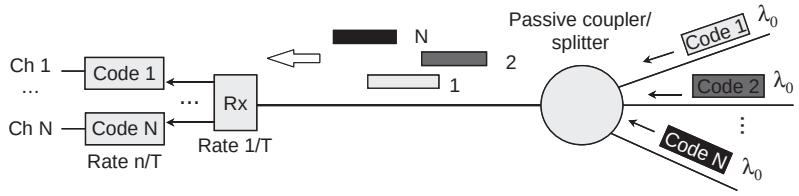


Figure 16.9 CDMA upstream.

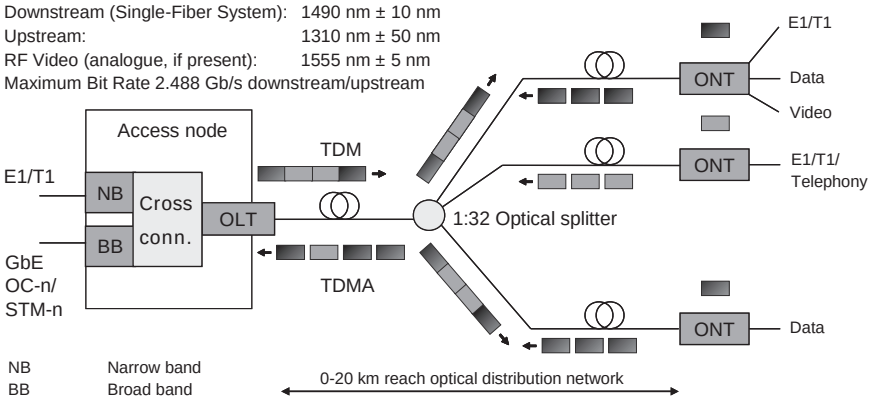


Figure 16.10 ITU FSAN PON access.

PONs need appropriate multiple-access schemes (TDMA in most cases) in order to share the common optical components between several customers. In addition, the relevant PONs (according to ITU-T G.983, G.984, and IEEE 802.3ah) only use *single* fibers between the CO (OLT) and the passive splitter/combiner, and between the splitter/combiner and any of the ONTs/ONUs. PONs must hence support single-fiber working (SFW) as indicated in Fig. 16.10.

SFW is usually done using WDM, that is, providing different wavelengths for the downstream and upstream directions, and separating/combining these at the Tx/Rx either with directional WDM or 3 dB couplers. In the ITU FSAN (full-service access network) standard, 1490 ± 10 nm is defined for the downstream, and 1310 ± 50 nm is used for the upstream. An additional 1555 ± 5 nm downstream wavelength can be used for video broadcast.

In Fig. 16.10, the (SONET/SDH) cross-connect may be replaced by a packet-oriented (Layer-2) switch or any other suitable aggregation device.

16.2. PON VARIANTS AND STANDARDS

Several PON approaches exist, which have been standardized in different standards. Generally, PONs are continuously evolving toward higher bandwidths, following customer demands for triple-play services with multi-channel video (VoD, IPTV) and high-speed Internet access. Today (2007), three main PON contenders can be identified:

- BPON (Broadband PON) according to ITU-T G.983.x, formerly referred to as APON (ATM PON)
- EPON (Ethernet PON) according to IEEE 802.3ah, also referred to as EFMP (Ethernet-in-the-First-Mile PON)
- GPON (Gigabit PON) according to the FSAN ITU standard G.984.x

Further (less relevant, at least as of 2007) variants include WDM PON and 10GPON.

The original PON standards focused on ATM as the Layer 2 protocol, evolving from work initiated at British Telecom. The standards were formalized in the full services access network (FSAN) consortium [2] and, eventually, in the International Telecommunications Union's ITU-T as the G.983.x specification. This work, now commonly known as APON or BPON, supports the majority of the early deployments around the world.

16.2.1. ATM PON (APON) and Broadband PON (BPON)

APON uses the ATM protocol, where the ATM header is complemented by an APON overhead. APON is still an efficient approach to establishing an optical access network other than P2P or ring-based fiber architectures.

Two APON versions have been defined:

- Symmetric: 155.52 Mb/s Downstream (DS) + Upstream (US)
- Asymmetric: 622.08 Mb/s DS + 155.52 Mb/s US

In the broadcast or downstream (DS) direction, a continuous ATM stream at a bit rate of 155 or 622 Mb/s is used. Dedicated physical layer OAM (PLOAM) cells are inserted into the datastream. The DS frame consists of 56 ATM cells for the basic rate of 155 Mb/s, scaling up to 224 cells for 622 Mb/s. At basic rate, 2 PLOAM cells are inserted (one at beginning, and one in middle); the other 54 cells are data ATM cells. The PLOAM cells contain grants for upstream transmission as well as OAM + P messages.

The upstream (US) is TDMA-based. For collision-free TDMA, the OLT sends permission to send data to the ONUs by sending grants via DS PLOAM cells. This is done in the form of bursts of ATM cells, with a 3-byte physical overhead

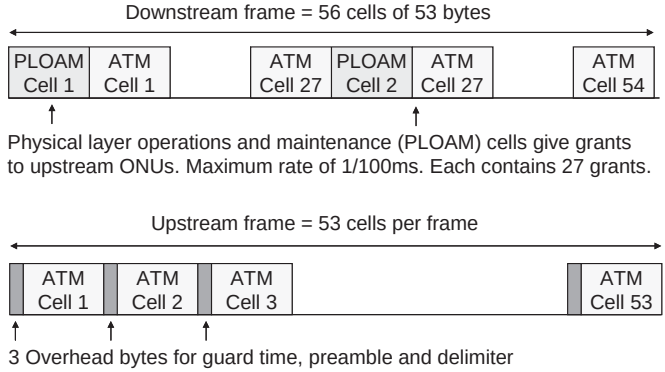

Downstream frame = 56 cells of 53 bytes
PLOAM Cell 1
ATM Cell 1
ATM Cell 27
PLOAM Cell 2
ATM Cell 27
ATM Cell 54
Physical layer operations and maintenance (PLOAM) cells give grants to upstream ONUs. Maximum rate of 1/100ms. Each contains 27 grants.
Upstream frame = 53 cells per frame
ATM Cell 1
ATM Cell 2
ATM Cell 3
ATM Cell 53
3 Overhead bytes for guard time, preamble and delimiter

Figure 16.11 G.983 APON/BPON frame format.

appended to each 53-byte cell. The US frame consists of 53 cells of 56 bytes each for the basic rate of 155 Mb/s. The APON frame format is shown in Fig. 16.11.

US transmission consists of either a data cell, containing ATM data in the form of VPs/VCs (virtual paths/circuits), or it may contain a PLOAM cell instead when granted a PLOAM opportunity from the central OLT.

Bidirectional communication in APON is using WDM with two wavelengths and SFW. Upstream is using 1310 ± 50 nm, whereas downstream is using 1490 ± 10 nm. In addition, in ITU-T G.983.3 a wavelength of 1555 ± 5 nm has been standardized for video broadcast overlay. This overlay offloads the line rate on the digital (DS) side and avoids requiring set-top boxes (STBs) on all TVs. Transport of 135 6-MHz RF channels ranging from 50 MHz to 870 MHz is possible.

Over time, the term APON led users to believe only ATM could be provided, so FSAN decided to broaden the name to *Broadband PON* (BPON), and also to further develop the respective functionality.

BPON is now standardized for symmetrical bit rates of 622 Mb/s and asymmetrical bit rates of 1244/622 Mb/s (DS/US), respectively. It supports native ATM, TDM (T1/E1) by circuit emulation, and Ethernet by emulation.

BPON enables splitting ratios of 1 : 32 (some systems 1 : 64). The total loss budget (including the infrastructure) is 28 dB, which enables maximum reach in excess of 20 km.

BPON is the predominant system in the United States at present. A complete BPON system with video overlay is shown in Fig. 16.12.

The system shown in Fig. 16.12 also contains the DCN (data communications network) for connecting the active PON components to a centralized management system (EMS/NMS, element/network management system, and OSS, operations support system, according to ITU-T M.3010).

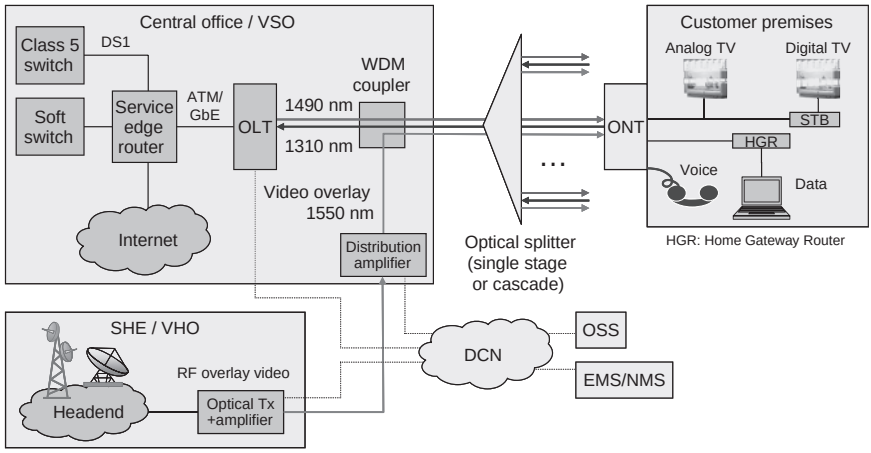


Figure 16.12 BPON with video overlay at 1550 nm acc. to FSAN/ITU-T G.983. VHO: Video Hub Office, SHE: Super Headend.

16.2.2. Gigabit PON (GPON)

APON/BPONs can have difficulties in scaling as carriers deploy more high-speed data and switched digital video services. APON/BPONs generally scale to 622 Mb/s downstream (CO to CP) and 155 Mb/s upstream, and, with ATM protocol overheads subtracted, support up to 448 Mb/s of usable/saleable bandwidth in the downstream direction. With bandwidth requirements per subscriber always increasing, that limits the number of users that can be served per PON. One of the main benefits of PON is being able to share the costs of the trunk(s) from the CO to the premises. If the Return on Investment (RoI) for these trunks is constrained by the bandwidth available per subscriber, the overall business case for PON can get lost.

In 2001, the FSAN group initiated a new effort for standardizing PON networks operating at bit rates of >1 Gb/s. Apart from the need to support higher bit rates, the overall protocol has been opened for reconsideration, and the sought solution should be the most efficient in terms of support for multiple services, OAM + P functionality, and scalability.

As a result, a new solution has emerged into the optical access market place—GPON (Gigabit PON), offering high bit rate support while enabling transport of multiple services, specifically data and TDM, in native formats and at high efficiency.

As part of the GPON effort, a gigabit service requirements (GSR) document has been put in place based on the collected requirements from all member service providers, representing the leading RBOCs and ILECs of the world. New

standards covering the GSR as well as the physical medium are defined in ITU-T standards G.984.1 and G.984.2, whereas G.984.3 covers the protocols.

The main requirements from the ITU documents were:

- Full service support—including voice (TDM, both SONET and SDH), Ethernet (10/100BaseT), ATM, leased lines, and more.
- Physical reach of at least 20 km with a logical reach support within the protocol of 60 km.
- Support for various bit rate options using the same protocol, including symmetrical 622 Mb/s, symmetrical 1.25 Gb/s, 2.5/1.25 Gb/s DS/US, and more.
- Strong operation administration and maintenance (OAM + P) capabilities offering end-to-end service management.
- Security at the protocol level for downstream traffic, which is necessary owing to the multicast nature of PON.

GPON supports two methods of encapsulation: the ATM and GPON encapsulation method (GEM). The ATM method is an evolution of existing APON/BPON standards, and voice, video, and data traffic can be ATM-encapsulated at the CPs for transport back to the CO.

With GEM, all traffic is mapped using a variant of the SONET/SDH generic framing procedure (GFP). GEM supports native transport of voice, video, and data without an added ATM or IP encapsulation layer. In addition to support of up to 2.5 Gb/s per PON, the use of GEM means that only 4 to 5% of this bandwidth is used for overhead, leaving the rest for revenue.

With GEM, traffic is carried in its native format without the costs of additional protocols. In comparison, GPON offers savings in bandwidth overhead for encapsulation and guarantees end-to-end clocking and QoS.

GPON frames are fixed at 125 μs. They support ATM and GEM payload within the same frame. Frames consist of a physical control block downstream (PCBd) and payload. The PCBd is used for synchronization, a DS OAM channel, and US bandwidth mapping. Bandwidth mapping is based on the service level agreement (SLAs) and the dynamic bandwidth requirement report (DBR) from the ONUs/ONTs. Specific periods within the U.S. frame are allocated for transmission of traffic. For management of the network, the OLT measures the power received from individual ONUs, allocates IDs to ONUs, discovers new ONUs added to the network, and collects performance parameters from each ONU.

In the upstream, frames contain various overhead bytes, including the physical layer overhead upstream (PLOu), which is responsible for the synchronization for new transmitters, the PLOAMu (US OAM channel), and the DBR.

According to the ITU standard, the DS has an operating wavelength of 1480 to 1500 nm. Minimum launch power is −4 dBm, whereas maximum launch power is +1 dBm, and the minimum receiver sensitivity is −25 dBm. The upstream uses

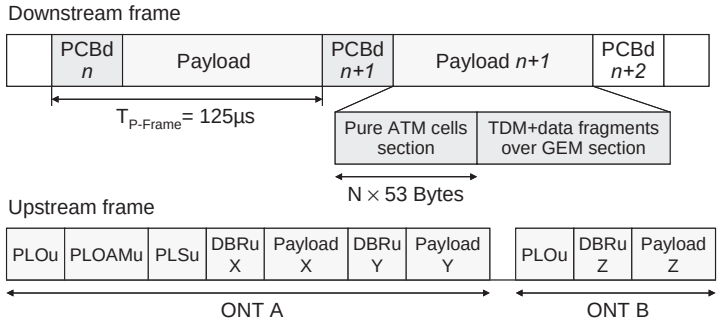


Figure 16.13 G.984 GPON frame format.

1260 to 1360 nm wavelength, minimum/maximum launch power is −3/+2 dBm, and the receiver sensitivity is −24 dBm, respectively. Again, an additional wavelength is considered for video overlay, which may be left unused due to the available downstream bandwidth.

GPONs are ideally suited for carrying both Ethernet/IP traffic and legacy voice and video services using a variant of GFP, which fully supports end-to-end clocking and other plesiochronous services. GPON also fully supports Ethernet signaling required for quality of service (QoS), class of service, and other advanced features. With GPON, PONs can evolve from carrying legacy TDM services to an all-IP/packet-based network. This allows carriers to make the decision about when to migrate to Voice-over-IP. The GPON frame structure is shown in Fig. 16.13.

With some implementations of GPON, the network can also be fully redundant either in a ring or tree architecture. With the use of CWDM, multiple GPONs can be carried over the same fiber, thereby increasing the RoI (Return on Invest) per capital expenditure on fiber deployment. The RoI can also be enhanced versus APON/BPON and EPON because GPON can support higher numbers of customers at the same per-customer bandwidth.

16.2.3. Ethernet PON (EPON, EFMP)

Ethernet for subscriber access networks, also referred to as Ethernet in the First Mile, or EFM, combine a minimal set of extensions to the IEEE 802.3 media access control (MAC) and MAC control sublayers with a family of Physical (PHY) layers. These physical layers include single-mode optical fiber and unshielded twisted pair (UTP) copper cable physical medium dependent sublayers (PMDs) for P2P connections in access networks.

The IEEE 802.3ah EFM standard also introduces the concept of Ethernet passive optical networks (EPONs), in which a point-to-multipoint (P2MP)

network topology is implemented with passive optical splitters, along with optical fiber PMDs that support this topology. In addition, a mechanism for network operations, administration, and maintenance (OAM) is included to facilitate network operation and troubleshooting. The EFM standard also defines active P2P (AON) variants for copper- (EFMC) and fiber-based (EFMF) Ethernet access.

EPONs (also known as EFMPs) are supported in the market by the Ethernet first mile alliance (EFMA), which became part of the Metro Ethernet Forum (MEF) [3].

EPONs enable IP-based P2MP connections using passive fiber infrastructure. Up- and downstream are controlled using the multipoint control protocol (MPCP). The US makes use of TDMA.

EPON was mainly motivated by the disadvantages of ATM. These include the facts that dropped cells invalidate entire IP datagrams, that ATM imposes a cell tax on variable-length IP packets, and that ATM in general did not live up to its promise of becoming an inexpensive technology.

EPON, on the other hand, provides an IP data-optimized access network, considering the fact that Ethernet is by far the most relevant protocol in the access. It provides EPON encapsulation of all data in Ethernet fames. The EPON layer stack is compared to APON and GPON in Fig. 16.14.

Single-mode fibers are used for EPON. Single-fiber working is enabled by using 1300 nm for the US and 1500 nm for the DS, respectively. Splitting ratios of 1 : 4 to 1 : 64 are supported (typical: 1 : 16). The maximum optical power budget is 20 dB, enabling maximum link lengths of 10 to 20 km.

EPON provides a symmetrical bit rate of 1.25 Gb/s for Ethernet transport only. In the DS, Ethernet frames transmitted (broadcast) by the OLT pass through the 1 : N passive splitter and reach each ONU (with own MAC addresses). This is similar to a shared-media network. Almost 50% of the available bandwidth is required for the protocol overhead, leaving only ~600 Mb/s for revenue use.

In the US, data frames from any ONU will only reach the OLT due to the directional properties of the passive splitter/combiner. This is similar to an Ethernet P2P architecture. However, EPON frames from different ONUs transmitted

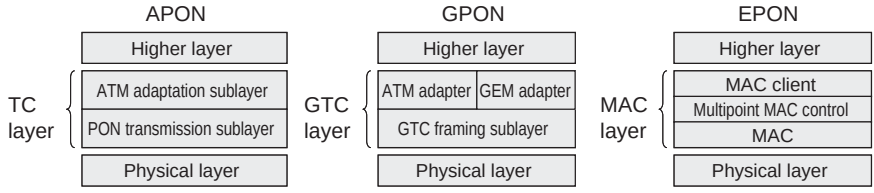


Figure 16.14 EPON layers compared to APON and GPON. (G)TC: (GPON) transmission convergence layer.

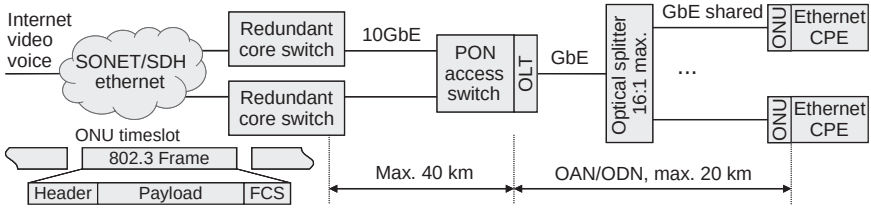


Figure 16.15 EPON (EFMP).

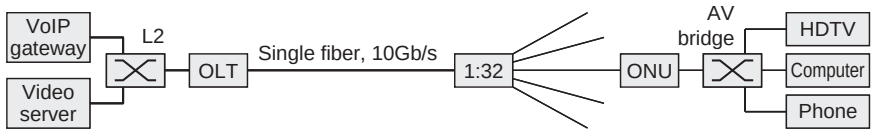


Figure 16.16 10 G EPON system.

simultaneously can still collide. Hence, ONUs need to share the trunk fiber channel capacity and resources.

An EPON access network is shown in Fig. 16.15.

The EPON system provides a very basic transport solution where cost-effective data-only services is the primary focus. EPONs are receiving a lot of attention in the Far East where missing pieces of the 802.3ah standard are being driven by NTT. There is not much interest in the United States.

EPON as a protocol is still under study within the IEEE EFM group.

16.2.4. 10G EPON

One of the major drawbacks of EPON is its limited bandwidth, which does not allow high per-ONU bandwidths in the range of 50 to 100 Mb/s, together with higher splitting ratios. (The same is true for BPON and, partly, GPON.) In IEEE 802.3av, a 10 G EPON is proposed in order to avoid this problem. This standard will mainly provide the PHY adaptation. It is still (2007) a living standard and has not been finalized.

Currently, both 10 G/1 G and 10 G/10 G DS/US bit rates are considered, which provides massively improved bandwidth for DS multicast (HDTV) services. The main differences in 1 G vs. 10 G US are related to available bandwidth, receiver sensitivity, and resulting cost.

In 802.3av, distances of 10 km and 20 km are discussed as well as splitting ratios of 1 : 16 and 1 : 32, respectively. The standard is based on SFW. A block diagram of a complete 10 G EPON system is shown in Fig. 16.16.

10 G EPON is capable of supporting 32 ONUs at high, dedicated (unicast) bit rates in excess of 100 Mb/s, even at low efficiencies of the underlying TDM/TDMA mechanism. It thus supports the high per-ONU bit rates that are commonly expected in the near future (starting 2010–2012). For higher splitting ratios, 10 G in PONs has to be complemented by WDM techniques.

16.2.5. WDM PON

The PONs discussed in Sections 16.2.1 to 16.2.4 impose some limitations due to the TDMA approach that is implemented and the use of passive 1 : N power splitters. These PONs may have bandwidth limitations in certain implementations, and the OLT and ONUs have to work at the aggregate bit rate. Capacity upgrades, or the addition of an ONU, are complicated with TDM/TDMA, and there are privacy/security issues due to the broadcast of DS information. Network integrity can be affected by a single ONU, which corrupts the entire upstream transmission. In addition, 1 : N power splitters lead to severe budget penalties, thus limiting maximum link lengths (example: a 1 : 32 splitter imposes >15 dB of addition loss).

The problems mentioned in this chapter can be solved with WDM PONs. Here, ONUs are assigned individual wavelengths. This provides higher bandwidth, in addition each ONU works on the individual bit rate rather than the aggregate (WDM) bit rate. Since ONUs are separated via physical wavelengths, privacy/security and network integrity aspects can be considered.

Alternatively, WDM PONs can be combined with any of the PONs described earlier, in particular EPON and GPON. This leads to combinations of WDM/WDMA and TDM/TDMA techniques and is the basis of considerations that lead to massively scalable PONs that support splitting ratios of up to 1 : 1000. Using WDM techniques—in particular amplification—these PONs can also support enhanced distances in the range of 100 km. This leads to the concept of active PONs, which play an important role in the considerations regarding the future metro access and backhaul convergence; see, for example [8].

So far, WDM PON has not been standardized. Many approaches are discussed today, which all have certain advantages and disadvantages. Following, an overview on some of the WDM PON proposals is given. For further reading, refer to [9].

The basic WDM PON architecture is shown in Fig. 16.17. It makes use of a fixed-wavelength laser array or a multifrequency laser (MFL). The *Broadcast+Select* architecture shown in Fig. 16.17 broadcasts all wavelengths in the DS through the passive splitter. Each ONU uses an individual wavelength for the US, which is combined in the passive combiner. Obviously, this still imposes the loss of the splitter and also the broadcast security issues. In addition, *no identical ONUs* can be used.

An alternative is the AWG-based (arrayed waveguide grating) wavelength-routing PON. Here, the passive splitter/combiner is replaced by an AWG wave-

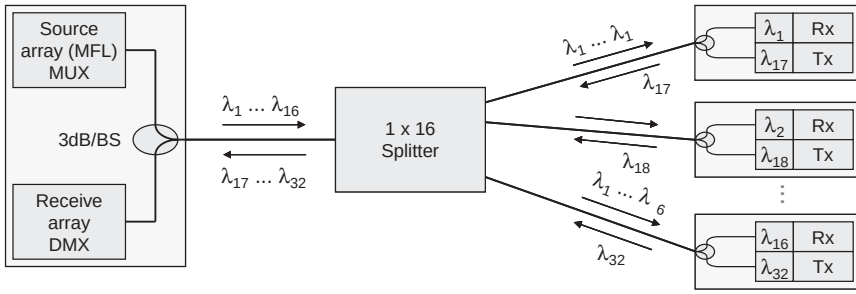


Figure 16.17 Basic WDM PON architecture with splitter/combiner in passive node (MFL: multi-frequency laser, BS: band splitter).

length router. This offers lower insertion loss (AWGs have typical insertion loss of 5 to 6 dB, independent from the number of wavelengths).

Through the periodic assignment of wavelengths to the output ports of the AWG, a certain degree of flexible wavelength and hence bandwidth assignment to the ONUs is possible in the wavelength-routing PON approach. This requires changing (tuning) the wavelengths in the OLT. In addition, no wavelength-selective (individual) *receivers* are necessary, thus simplifying the ONUs. However, different ONU Tx wavelengths are still necessary.

Complete ONU unification can be achieved by using a single wavelength for the US. The resulting WDM PON approach is also referred to as Composite PON (CPON) [9]. The CPON approach leads to identical ONUs, thus simplifying the WDM PON. This comes at the cost of total US bandwidth, and the need for TDMA techniques. The CPON approach uses a single wavelength for the US in order to unify the ONU design. Alternatively, ultra-broadband—LED—sources can be used in the ONUs. This leads to the local access router net (LARNET) approach [9]. In the AWG router, the broadband US signals are spectrally sliced according to Fig. 16.7. In the OLT, the LARNET uses circulators for SFW instead of 3 dB couplers or band splitters (BS). The LARNET approach leads to identical ONUs, but it suffers from poor power budget, which results from the use of LEDs and the related coupling efficiencies into single-mode fibers. Also, the US bandwidth is limited, and multiple-access (burst mode) techniques are required for the US.

Another approach that leads to unified ONUs makes use of broadband optical modulators, for example, semiconductor optical amplifiers (SOAs), together with the provisioning of so-called seed lasers. This approach was followed in the remote interrogation of terminal network (RITENET) [9]. In the RITENET ONUs, the receive signal (wavelength) is split and fed into the receiver and the (SOA) modulator, respectively. Hence, identical ONUs can be used, but DS and US use

the same wavelength, which leads to dual-fiber working. Together with the cost of SOA-based transmitters, this is the major disadvantage of the RITENET approach.

Many other WDM PON proposals have been made. These include the use of SFPs (small form-factor pluggables) for simplified ONU design, cascaded AWG filters for improved scalability, and EDFA amplification for enhanced PON distances and splitting ratios.

All these WDM PONs can be used in combination with other PON technologies, that is, GPON and EPON. They can also easily be combined with (DS or US) bit rates of 1 G, 2 G5, 4 G, or 10 G, respectively. Over time, this will lead to ultimately flexible, scalable, ultra-broadband access infrastructures. Today (2007), however, no significant deployments can be seen.

16.2.6. PON Deployment Reference

An early example of a massive PON rollout is the FiOS services, which is offered by Verizon Communications Inc. in some areas of the United States. FiOS is an FTTP telecommunications service. It stands for fiber-optic service. FiOS started as a pilot program in Keller, Texas, but availability of the service has now expanded to many states.

There are several tiers of residential Internet service. Availability depends on the location of the customer. Speeds range up to 50/10 Mb/s DS/US in certain market areas. Higher speeds are available in highly competitive areas, such as Greater Boston. In addition to residential offerings, FiOS business service is available, with higher upload speeds, static IP addresses, and no blocked ports.

FiOS TV started in Keller, Texas, in October 2005. At the end of March 2007, TV services were available to 3.1 million premises in 10 states. The service began with 293 channels of video and 1800 choices for video on demand. The TV services beyond the basic 25 channels require a digital set-top box or CableCard to receive and decode the television signal.

Verizon also offers analog services (POTS) over FiOS. While FiOS phone service offers digital audio quality compared to standard copper phone lines, power outages may affect service availability. Unlike standard phone lines, the FiOS service depends on power at the customer premises.

Optical fiber extends from central offices to unpowered hubs, in which the optical signal is optically split up to 32 ways. The active components adhere to the ITU-T G.983 standard, BPON:

- 622 Mb/s DS @ 1490 nm
- 155 Mb/s US @ 1310 nm
- RF video overlay @ 1550 nm

In 2007, GPON started replacing the BPON technology.

16.3. COMPARISON OF MAIN PON APPROACHES

Comparing the relevant (non-WDM) PON standards as discussed so far, we see that carriers primarily have to decide between GPON and EPON (with GPON being favored in the United States, and EPON in Asia-Pacific).

GPON is BPON's (and hence APON's) natural evolution. It provides several advantages over EPON:

- GPON offers higher line rates and greater bandwidth efficiency and flexibility for high-speed data services than EPON. This can lead to lower fiber and optics cost.
- With transport of (multimode) IP/Ethernet, ATM, and TDM payloads, GPON supports network upgrades for the large base of legacy services as well as for emerging IP triple play services, including TV overlay.
- Support for longer reach from the CO to the user.

These characteristics make GPON an appropriate choice if legacy services have to be supported. On the other hand, EPON—despite its lower efficiency—*natively* supports Ethernet (GbE) as the most relevant access protocol in the market. Advantages of EPON include:

- EPON may, over time, lead to cheapest (GbE) interface cost.
- EPON is fully integrated into the EFM–OAM approach. As such, it can consistently be combined with an AON approach (active Ethernet P2P and Ethernet P2MP), allowing mix + match combinations of PONs and AONs under one homogenous (Ethernet) management system.

For further comparison, Table 16.1 lists relevant parameters for PONs.

Table 16.1

Parameters of Relevant PONs.

	Standard	DS [Mb/s]	US [Mb/s]	Efficiency for Split of 10/90 TDM/Data	Efficiency for Split of 20/80 TDM/Data
APON	ITU-T FSAN	155 622	155 155	71%	72%
BPON	G.983 FSAN	155 622	155 622	71%	72%
EPON	IEEE 802.3ah	10–1000 (1250)	10–1000 (1250)	55%	55%
GPON	G.984 FSAN	1244 2488	155/622 1244 2488	93%	94%

The comparison does not include WDM PONs. This approach can, however, be combined with all PONs listed in Table 16.1, for example, to overcome scaling problems. The resulting WDM PONs can then make use of the GPON or EPON/EFM OAM capabilities, together with the respective multiple-access techniques.

REFERENCES

1. de Albuquerque, A. A., et al. 2004, February. Field trials for fiber access in the EC. *IEEE Comm. Magazine*, pp. 40–48.
2. www.fsanweb.org
3. www.metroethernetforum.org
4. ITU-T G.983.
5. ITU-T G.984.
6. IEEE 802.3ah, EFM.
7. IEEE 802.3av, 10GEPON.
8. Davey, R., et al. 2007, September. Long-reach access and future broadband network economics. ECOC ′07, Berlin.
9. Banerjee, A., et al. 2005, November. Wavelength-division multiplexed passive optical network (WDM-PON) technologies for broadband access: a review [Invited]. *Journal of Optical Networking* 4, no. 11:737–758.

Part III

Applications & Industry Standards

This page intentionally left blank

17

Optical Interconnects for Clustered Computing Architectures

David B. Sher
Mathematics/Statistics/CMP Department, Nassau Community College, Garden City, NY

Casimer DeCusatis
IBM Corporation, Poughkeepsie, NY

17.1. INTRODUCTION

The advantages of optical communications are well known for communicating between computing systems or other large-scale communication applications like telephony and cable television. However, using fiber optics for communicating within a single computer, or between the elements of a clustered computer architecture, is a relatively new and emerging field. While early attempts were expensive and inefficient because of the conversion costs between electronic and optical signals, improvements in optical interconnect technology have made these solutions practical over distances as short as a few hundred meters or less.

This chapter focuses on applications for optical communication within a single computing system or a computing cluster, sometimes referred to as a symmetric multiprocssor (SMP) network. The boundaries of such networks are not clearly defined, and the optical interconnects are not standardized with the same rigor as other local area or metropolitan area networks. One possible configuration is shown in Fig. 17.1, which illustrates the relationship between a system area network and other components of the computer architecture. Using this approach rather than a system bus is more beneficial on larger computers, particularly multiprocessors, though a single processor that serves a large number of fast IO sources may also benefit. Thus, the earliest applications for optical interconnect have been within supercomputers and mainframes. While many commercial systems continue to use copper interconnects, several are now adopting optics for both intraprocessor communication in cases where large numbers of processors are required and for connecting processors to resources like memory and I/O.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

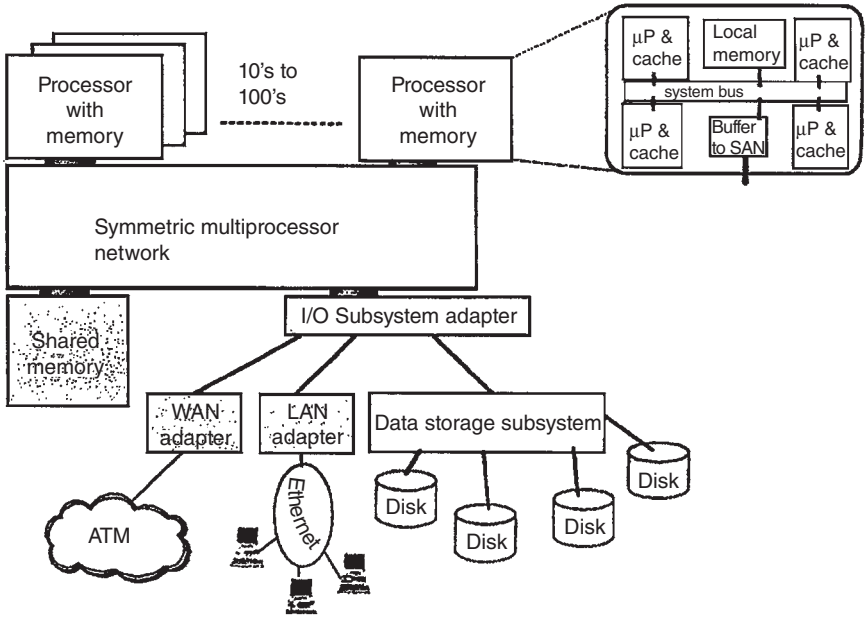


Figure 17.1 Diagram of symmetric multiprocessor network, illustrating how it interfaces to other data communications networks.

Although system area networks are not yet standardized, they tend to share the following common characteristics (performance values are current as of this writing, and although the absolute numbers will improve over time it appears that the distinctions between other optical networks will remain).

(a) Bandwidth
The bandwidth requirements of optical interconnects are higher than those of a LAN or other networks, ideally as close as possible to the bandwidth of a system bus. Practically, this will mean a bidirectional bus with bandwidth exceeding 10 to 50 Mbit/s in both directions.

(b) Latency
Because communications are intended to approach the performance over a system bus, where latency is measured in hundreds of ns, the total latency across an optical interconnect should be less than 1 microsecond, preferably as low as 10–100 ps.

(c) Scalability
Optical interconnects must support economic scalability from a few computing nodes up to hundreds or thousands of nodes or more. There may be practical issues with programming applications scalable to this

range, although some SIMD machines could be considered to be in this range.

17.2. COPPER *VS.* OPTICAL TECHNOLOGY TRADEOFFS

Computer communications can be viewed as a hierarchy of interconnects:

1. Intrachip connections between transistors, gates, and functional blocks
2. Interchip connections, ranging from direct chip to chip on a single-block circuit board to bused connections between multiple chips on multiple circuit board machines within a single machine room
3. Rack-to-rack connections, for closely packed racks in large machines within a single machine room
4. Buildingwide or campuswide connections over a LAN
5. Intercampus connections across a MAN
6. Wide area network (WAN) machine connections with a large city or between cities

The general characteristics of these types of interconnects are shown in Table 17.1. The actual numbers are approximate and will not, of course, be equivalent for a personal computer and a supercomputer. However, with the increasing convergence between long-end and high-end computer technology, the values are correct over a perhaps surprisingly wide range of computing systems. The basic electronics that drive each of these levels is approximately the same. To within a close approximation, the speed of the transistor that drives a wire to the other side of a very large-scale integration (VLSI) chip is the same as the speed of the transistor that drives a coaxial line or a laser driver.

A number of factors determine whether a given level of interconnect is implemented in electronic or optical transmission technology. Optical data transmission

Table 17.1

Hierarchy of Computer System Interconnect Technologies.

Interconnect Level	Technology Options	Bit Rate Per Line	Signal Lines Per Connection	Line Distance
Intrachip	Metal on CMOS	50–1000 Mb/s	64–256	<25 mm
Chip to chip	Metal on circuit board	30–1000 Mb/s	32–128	1–100 cm
Rack to rack	Twisted pair, coax, multimode fiber	20–200 Mb/s	16–54	1–10 m
LAN	Twisted pair, coax, multimode fiber	10–100 Mb/s	1–10	10–1000 m
MAN	Single-mode fiber, multimode fiber	56 Kb/s-1.544 Mb/s	1–10	1–10 km
WAN	Single-mode fiber	2488 Mb/s	1	>10 km

technology is, in general, more expensive than electronic data transmission using a simple cost-per-bit metric. For the longest distances (multi-km), the cabling and repeater costs of copper outweigh the optical module costs, and optical technology becomes cost-effective. At shorter distances, other considerations such as bandwidth required by the application, thermal dissipation, and electrical power consumption can make optical interconnect the preferred choice. Independent of technology, the cost of an interconnect increases dramatically with distance—for example, a single interchip wire costs less than a penny, whereas a LAN cable costs tens to hundreds of dollars and a WAN line costs thousands of dollars. Copper wire is more cost-sensitive to a figure of merit that includes the product of data rate and distance figure of merit. Technical considerations such as signal distortion, electromagnetic interference, cost and weight of cables, and crosstalk are also more significant for electronic data transmission than they are for optical links.

There are currently no industry standards that specify the properties of an intramachine optical communication link. Many early adopters have implemented some version of a switched interconnect fabric as shown in Fig. 17.2; these include InfiniBand (see Chapter 24), Fibre Channel (see Chapter 20), or low-latency Ethernet (see Chapter 22). Some examples are shown in Table 17.2. Many storage networks use small computer system interconnect (SCSI) protocols for commu-

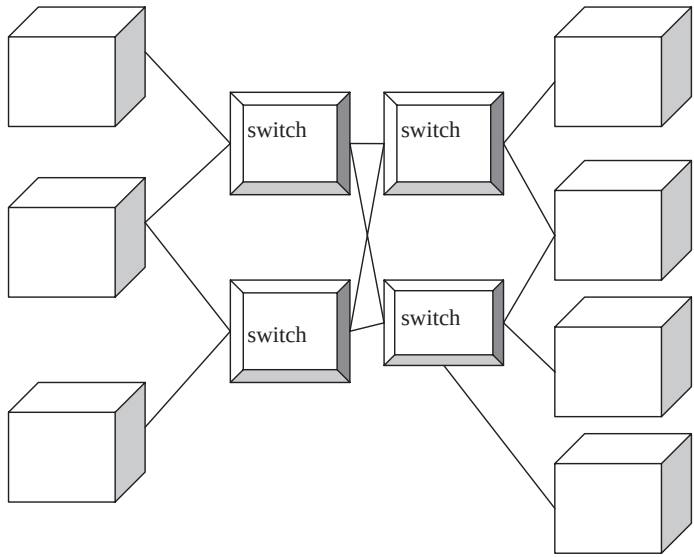


Figure 17.2 Block diagram of switched interconnect fabric.

Table 17.2
Examples of Switched Fabric Interconnects.

Standard	Properties	Suppliers of Optical Fiber Implementations in 2007
Fibre Channel	Currently the most common. Comes in 1 Gbit/s, 2 Gbit/s, 4 Gbit/s, and 8 Gbit/s variants (10 Gbit/s is under development)	Apple, Qlogic, CS Electronics, Hewlett Packard, ADVA Optical Networking, Emulex, Emcore, TMC, LSI, Cisco, SGI, Dell, Avid, and Sun.
iSCSI	SCSI over TCP/IP	Adaptec, Qlogic, Cisco, Agilent Technologies, IBM, and Sunstar
InfiniBand (IB)	SCSI over IB and/or TCP/IP over IB	Emcore, Zarlink, Amphenol, Intel, Alvesta, GE, IBM?
Myrinet	For high-performance clusters, all optical. Low.	Myricom
ATA over Ethernet	ATA over Ethernet	Coraid supplies standard (ordinary ethernet is used)
HyperSCSI	SCSI over Ethernet	No suppliers

nication between servers and storage devices, though instead of its low-level interface they use some type of mapping layer. These protocol standards are intended to be modular and independent of the physical media (copper or optics). However, we note that interswitch links and other features on many switches do not interoperate between switches from different manufacturers. Indeed, most intramachine applications do not require computers from different manufacturers to interoperate, making it possible to differentiate systems based on their interconnect networks. As optics becomes more tightly integrated within the computer architecture, new types of transceiver and cable packaging are expected to emerge. Higher volume implementations may lead to ad hoc industry standards for the building blocks in an optical interconnect system.

17.3. TOPOLOGIES

Historically, many interconnect topologies have been demonstrated for intramachine communications, including pipelines, trees, and cube-connected cycles. However, three topologies as of this writing are as follows.

1. Full connectivity using a crossbar or bus. The historic C.mmp processor used a crossbar to connect the processors to memory. Computers with small numbers of processors (like a typical parallel sysplex system or tandem system) can use this topology, but it becomes cumbersome with large (more than 16) processors since every processor must be able to simultaneously directly communicate with each other. This topology requires a fan in and

fan out proportional to the number of processors making large networks difficult.

2. Torus and Allied topologies where an N processor machine requires $\sqrt{N}$ processors to relay messages. For example, the Goodyear massively parallel processor (MPP) machine was laid out as a torus. Each processor in a torus is connected to four neighbors (north, south, east, west). The most western processors are connected to the most eastern ones, and the most northern processors are connected to the most southern ones.

   Such topologies are easy to lay out on silicon, so multiple processors can be placed on a single chip and many such chips can be easily placed on a board. Such technology may be particularly appropriate for computations that are spatially organized. This topology also has constant fan in and fan out.

   Many of the top supercomputers are connected in a three-dimensional torus. This is similar to the classic torus except instead of each processor connecting to four neighbors (north, south, east, west), they are connected to 6 neighbors (north, south, east, west, up, down).
3. Switched fabric: Switched fabric simulates full connectivity without actually directly connecting every component. It takes advantage of the fact that not every component will be simultaneously accessing every other one, and uses redundant paths through the switches to make collisions unlikely.

17.4. LARGE-SCALE COMPUTING

Of the top 10 supercomputers in the world (as of June 2007 on the Web site http://www.top500.org) the first 7 used a 3-D torus topology for interconnections. The last 3 all used various versions of a switched fabric. Switched fabric simulates total interconnectivity but allows for high bandwidth and low latency. In particular, the Infiniband network used by the seventh most powerful supercomputer was actually designed for micro-and minicomputer systems. Currently, it is applied primarily to clusters. Infiniband is of particular relevance because it has been implemented with both optical fiber and copper. Also 26 of the top 100 supercomputers use Infiniband as their network interconnect.

Myrinet is an important switched fiber network that is designed specifically for high-performance clusters. It is a fiber-optic-based system, though some of the implementations also accept copper cables. Nine of the top 100 supercomputers use Myrinet as their network interconnect.

Hence, currently the most important network topologies are the classic bus that offers timeshared total interconnectivity, the switched fabric that simulates total interconnectivity though a network of switches, and for high-performance computationally intensive applications the 3-D torus. Table 17.3 lists some of the

Table 17.3
The World's Fastest Supercomputers, as of Mid-2007.

Computer	Interconnect	Web Site
Blue Gene/L	3-D Torus	http://www.llnl.gov/asc/computing_resources/bluegenel/bluegene_home.html
Jaguar—Cray XT4/XT3	3-D Torus	http://info.nccs.gov/resources/jaguar
Cray Red Storm	3-D Torus	http://www.cs.sandia.gov/platforms/RedStorm.html
Various Blue Gene Solutions	3-D Torus	http://www.newyorkblue.bnl.gov/ http://www.rpi.edu/research/ccni/
ASC Purple	3-D Torus	http://www.llnl.gov/asc/computing_resources/purple/purple_index.html
Abe PowerEdge	Infiniband (switched fabric)	http://www.ncsa.uiuc.edu/
MareNostrum BladeCenter JS21	Myrinet (switched fabric)	http://www.bsc.es/
HLRB II	Numalink (SGI switched fabric bus)	http://www.lrz-muenchen.de/services/compute/hlrb/
Thunderbird	Infiniband (Switched fabric)	http://www.cs.sandia.gov/platforms/Thunderbird.html
Tera10	QsNet (fat tree)	http://fr.wikipedia.org/wiki/TERA-10
Columbia	Infiniband (switched fabric)	http://www.nas.nasa.gov/About/Projects/Columbia/columbia.html
Tsubame	Infiniband (switched fabric)	http://www.gsic.titech.ac.jp/
Lonestar	Infiniband (switched fabric)	http://www.tacc.utexas.edu/resources/hpcsystems/#lonestar

fastest computers in 2007, their interconnect between the processors and memory, and a Web site where more details about the system are available.

17.5. PARALLEL SYSPLEX AND GDPS

High-end computer systems running over MANs are proving to be a near-term application for multi-terabit communication networks. Large computer systems require dedicated storage area networks (SANs) to interconnect with various types of direct attach storage devices (DASD), including magnetic disk and tape, optical storage devices, printers, and other equipment. This has led to the emergence of client-servers-based networks employing either circuit or packet switching, and the development of network-centric computing models. In this approach, a high-bandwidth, open-protocol network is the most critical resource in a computer system, surpassing even the processor speed in its importance to overall performance. The recent trend toward clustered, parallel computer architectures to enhance performance has also driven the requirement for high-bandwidth fiber-optic coupling links between computers. For example, large water-cooled mainframe computers using bipolar silicon processors are being replaced by smaller,

air-cooled servers using complementary metal oxide semiconductor (CMOS) processors. The performance of these new processors can far surpass that of older systems because of their ability to couple together many central processing units in parallel. One widely adopted architecture for clustered mainframe computing is known as a geographically dispersed parallel sysplex (GDPS). In this section, we will describe the basic features of a GDPS and show how this architecture is helping to drive the need for high-bandwidth dense wavelength division multiplexing (DWDM) networks.

In 1994, IBM announced the Parallel Sysplex architecture for the System/390 mainframe computer platform (note that the S/390 has recently been rebranded as the IBM eServer z series). This architecture uses high-speed fiber-optic data links to couple processors together in parallel [1–4], thereby increasing capacity and scalability. Processors are interconnected via a coupling facility, which provides data caching, locking, and queuing services; it may be implemented as a logical partition rather than a separate physical device. The gigabit links, known as InterSystem Channel (ISC), HiPerLinks, or Coupling Links, use long-wavelength (1300-nm) lasers and single-mode fiber to operate at distances up to 10 km with a 7-dB link budget (HiPerLinks were originally announced with a maximum distance of 3 km, which was increased to 10 km in May 1998). If good quality fiber is used, the link budget of these channels allows the maximum distance to be increased to 20 km. When HiPerLinks were originally announced, an optional interface at 531 Mbit/s was offered using short-wavelength lasers on MM fiber. The 531 Mbit/s HiPerLinks were discontinued in May 1998 for the G5 server and its follow-ons. A feature is available to accommodate operation of 1 Gbit/s HiPerLinks adapters on multimode fiber, using a mode conditioning jumper cable at restricted distances (550 meters maximum).

The physical layer design is similar to the ANSI Fibre Channel Standard, operating at a data rate of 1.0625 Gbit/s, except for the use of open fiber control (OFC) laser safety on long-wavelength (1300-nm) laser links (higher order protocols for ISC links are currently IBM proprietary). Open fiber control is a safety interlock implemented in the transceiver hardware; a pair of transceivers connected by a point-to-point link must perform a handshake sequence in order to initialize the link before data transmission occurs. Only after this handshake is complete will the lasers turn on at full optical power. If the link is opened for any reason (such as a broken fiber or unplugged connector), the link detects this and automatically deactivates the lasers on both ends to prevent exposure to hazardous optical power levels. When the link is closed again, the hardware automatically detects this condition and reestablishes the link. The HiPerLinks use OFC timing corresponding to a 266-Mbit/s link in the ANSI standard, which allows for longer distances at the higher data rate. Propagating OFC signals over DWDM or optical repeaters is a formidable technical problem, which has limited the availability of optical repeaters for HiPerLinks. OFC was initially used as a laser eye safety

feature; subsequent changes to the international laser safety standards have made this unnecessary, and it has been discontinued on the most recent version of z series servers. The 1.06-Gbit/s HiPerLinks will continue to support OFC in order to interoperate with installed equipment; this is called "compatibility mode." There is also a 2.1 Gbit/s HiPerLink channel, also known as ISC-3, which does not use OFC; this is called "peer mode."

Since all the processors in a GDPS must operate synchronously, they all require a multimode fiber link to a common reference clock known as a Sysplex Timer (IBM model 9037). The sysplex timer provides a time of day clock signal to all processors in a sysplex; this is called an external timing reference (ETR). The ETR uses the same physical layer as an ESCON link, except that the data rate is 8 Mbit/s. The higher level ETR protocol is currently proprietary to IBM. The timer is a critical component of the Parallel Sysplex; the sysplex will continue to run with degraded performance if a processor fails, but failure of the ETR will disable the entire sysplex. For this reason, it is highly recommended that two redundant timers be used, so that if one fails the other can continue uninterrupted operation of the sysplex. For this to occur, the two timers must also be synchronized with each other; this is accomplished by connecting them with two separate, redundant fiber links called the control link oscillator (CLO). Physically, the CLO link is the same as an ETR link except that it carries timing information to keep the pair of timers synchronized. Note that because the two sysplex timers are synchronized with each other, it is possible that some processors in a sysplex can run from one ETR while others run from the second ETR. In other words, the two timers may both be in use simultaneously running different processors in the sysplex, rather than one timer sitting idle as a backup in case the first timer fails.

There are three possible configurations for a Parallel Sysplex. First, the entire sysplex may reside in a single physical location, within one data center. Second, the sysplex can be extended over multiple locations with remote fiber-optic data links. Finally, a multisite sysplex in which all data is remote copied from one location to another is known as a Geographically Dispersed Parallel Sysplex, or GDPS. The GDPS also provides the ability to manage remote copy configurations, automates both planned and unplanned system reconfigurations, and provides rapid failure recovery from a single point of control. There are different configuration options for a GDPS. The single-site workload configuration is intended for those enterprises that have production workload in one location (site A) and discretionary workload (system test platforms, application development, etc.) in another location (site B). In the event of a system failure, unplanned site failure, or planned workload shift, the discretionary workload in site B will be terminated to provide processing resources for the production work from site A (the resources are acquired from site B to prepare this environment, and the critical workload is restarted). The multiple-site workload configuration is intended

for those enterprises that have production and discretionary workload in both site A and site B. In this case, discretionary workload from either site may be terminated to provide processing resources for the production workload from the other site in the event of a planned or unplanned system disruption or site failure.

Multisite Parallel Sysplex or GDPS configurations may require many links (ESCON, HiPerLinks, and Sysplex Timer) at extended distances; an efficient way to realize this is to use wavelength-division multiplexing technology. Multiplexing wavelengths is a way to take advantage of the high bandwidth of fiber-optic cables without requiring extremely high modulation rates at the transceiver. This type of product is a cost-effective way to utilize leased fiber-optic lines, which are not readily available everywhere and may be very high cost (typically, the cost of leased fiber (sometimes known as dark fiber) where available is $300/mile/month). Traditionally, optical wavelength-division multiplexing (WDM) has been widely used in telecom applications, but has found limited usage in datacom applications. This is changing, and a number of companies are now offering multiplexing alternatives to datacom networks that need to make more efficient use of their existing bandwidth. This technology may even be the first step toward development of all-optical networks. For Parallel Sysplex applications, the only currently available WDM channel extender that supports GDPS (Sysplex Timer and HiPerLinks) in addition to ESCON channels is the IBM 2029 Fiber Saver (5–8) as described in Chapter 15 (note that the 9729 optical wavelength-division multiplexer also supported GDPS but has been discontinued; other DWDM products are expected to support GDPS in the future, including offerings from Nortel and Cisco). The 2029 allows up to 32 independent wavelengths (channels) to be combined over one pair of optical fibers, and to extend the link distance up to 50-km point-to-point or 35 km in ring topologies. Longer distances may be achievable from the DWDM using cascaded networks or optical amplifiers, but currently a GDPS is limited to a maximum distance of 40 km by timing considerations on the ETR and CLO links (the Sysplex Timer documents support for distances up to only 26 km, and the extension to 40 km requires a special request from IBM via RPQ 8P1955).

These timing requirements also make it impractical to use time division multiplexing (TDM) or digital wrappers in combination with DWDM to run ETR and CLO links at extended distances; this implies that at least 4 dedicated wavelengths must be allocated for the Sysplex Timer functions. Also note that since the Sysplex Timer assumes that the latency of the transmit and receive sides of a duplex ETR and CLO link are approximatly equal, the length of these link segments should be within 50 m of each other. For this reason, unidirectional 1 + 1 protection switching is not supported for DWDM systems using the 2029; only bidirectional protection switching will work properly. Even so, most protection schemes cannot switch fast enough to avoid interrupting the Sysplex Timer and HiPerLinks operation. HiPerLinks in compatibility mode will be interrupted by

their open fiber control, which then takes up to 10 seconds to reestablish the links. Timer channels will also experience loss of light disruptions, as will ESCON and other types of links. Even when all the links are reestablished, the application will have been interrupted or disabled and any jobs that had been running on the sysplex will have to be restarted or reinitiated, either manually or by the host's automatic recovery mechanisms depending on the state of the job when the links were broken. It is therefore recommended that continuous availability of the applications cannot be insured without using dual-redundant ETR, CLO, and HiPerLinks. Protection switching merely restores the fiber capacity more quickly; it does not ensure continuous operation of the sysplex in the event of a fiber break.

To illustrate the use of DWDM in this environment, consider the construction of a GDPS between two remote locations for disaster recovery as shown in Fig. 17.3. There are four building blocks for a Parallel Sysplex; the host processor (or parallel enterprise server), the coupling facility, the ETR (Sysplex Timer), and disk storage. Many different processors may be interconnected through the coupling facility, which allows them to communicate with each other and with data stored locally. The coupling facility provides data caching, locking, and queuing (message passing) services. By adding more processors to the configuration, the overall processing power of the sysplex (measured in millions of instructions per second or MIPS) will increase. It is also possible to upgrade to more powerful processors by simply connecting them into the sysplex via the coupling facility. Special software allows the sysplex to break down large database applications into smaller ones, which can then be processed separately; the results are combined to arrive at the final query response. The coupling facility may be implemented as either a separate piece of hardware or a logical partition of a larger system. The HiPerLinks are used to connect a processor with a coupling facility. Since the operation of a Parallel Sysplex depends on these links, it is highly recommended that redundant links and coupling facilities be used for continuous availability.

Thus, in order to build a GDPS, we require at least one processor, coupling facility, ETR, and disk storage at both the primary and secondary locations, which we will call site A and site B. Recall that one processor may be logically partitioned into many different sysplex system images; the number of system images determines the required number of HiPerLinks. The sysplex system images at site A must have HiPerLinks to the coupling facilities at both site A and B. Similarly, the sysplex system images at site B must have HiPerLinks to the coupling facilities at both site A and B. In this way, failure of one coupling facility or one system image allows the rest of the sysplex to continue uninterrupted operation. A minimum of two links are recommended between each system image and coupling facility. Assuming there are S sysplex system images running on P processors and C coupling facilities in the GDPS, spread equally between site A and site B, the total number of HiPerLinks required is given by

$$\text{\# HiPerLinks} = S * C * 2 \tag{17.1}$$

In a GDPS, the total number of intersite HiPerLinks is given by

$$\text{intersite \# HiPerLinks} = S * C \tag{17.2}$$

The Sysplex Timer (9037) at site A must have links to the processors at both site A and B. Similarly, the 9037 at site B must have links to the processors at both site A and B. There must also be two CLO links between the timers at sites A and B. This makes a minimum of 4 duplex intersite links, or 8 optical fibers without multiplexing. For practical purposes, there should never be a single point of failure in the sysplex implementation; if all the fibers are routed through the same physical path, there is a possibility that a disaster on this path would disrupt operations. For this reason, it is highly recommended that dual physical paths be used for all local and intersite fiber optic links, including HiPerLinks, ESCON, ETR, and CLO links. If there are P processors spread evenly between site A and site B, then the minimum number of ETR links required is given by

$$\text{\# ETR links} = (P * 2) + 2 \text{ CLO links} \tag{17.3}$$

In a GDPS, the number of intersite ETR links is given by

$$\text{intersite \# ETR links} = P + 2 \text{ CLO links} \tag{17.4}$$

These formulas are valid for CMOS-based hosts only; note that the number of ETR links doubles for ES/9000 multiprocessor models due to differences in the server architecture.

In addition, other types of intersite links such as ESCON channels allow data access at both locations. In a GDPS with a total of N storage subsystems (also known as direct access storage devices, or DASD), it is recommended that there be at least 4 or more paths from each processor to each storage control unit (based on the use of ESCON directors at each site). Thus, the number of intersite links is given by

$$\text{intersite \# storage (ESCON) links} = N * 4 \tag{17.5}$$

In addition, the sysplex requires direct connections between systems for cross-system coupling facility (XCF) communication. These connections may be provided by either ESCON channel-to-channel links or HiPerLinks. If coupling links are used for XCF signaling, then no additional HiPerLinks are required beyond those given by Eqs. (17.1) and (17.2). If ESCON links are used for XCF signaling, at least two inbound and two outbound links between each system are required, in addition to the ESCON links for data storage discussed previously. The minimum number of channel-to-channel (CTC) ESCON links is given by

$$\text{\# CTC links} = S * (S - 1) * 2 \tag{17.6}$$

For a GDPS with S_A sysplex systems at site A and S_B sysplex systems at site B, the minimum number of intersite channel-to-channel links is given by

$$\text{intersite \# CTC links} = S_A * S_B * 4 \tag{17.7}$$

Since some processors also have direct LAN connectivity, it may be desirable to run some additional intersite links for remote LAN operation as well.

As an example of applying these equations, consider a GDPS consisting of two system images executing on the same processor and a coupling facility at site A, and the same configuration at site B. Each site also contains one primary and one secondary DASD subsystem. Sysplex connectivity for XCF signaling is provided by ESCON CTC links, and all GDPS recommendations for dual redundancy and continuous availability in the event of a single failure have been implemented. From Eq. (17.7), the total number of intersite links required is given by

$$\begin{aligned}
&\text{\# of intersite links:}\\
&\text{\# CTC links} = S_A * S_B * 4 = 2 * 2 * 4 = 16\\
&\text{\# timer links} = P + 2 = 2 + 2 = 4\\
&\text{\# HiPerLinks} = S * C = 4 * 2 = 8\\
&\text{\# storage (DASD) links} = N * 4 = 8 * 4 = 32
\end{aligned} \tag{17.8}$$

or a total of 60 intersite links. These expressions do not apply to enhancements such as STP links (see following section); such installations must be treated on a case-by-case basis. Either ESCON or Fibre Channel links may be used for the direct connection between local and remote DASD via the peer-to-peer remote copy (PPRC) protocols. Other types of storage protocols may be used for the DASD connections. In the future, new protocols for clustering may be introduced as a replacement for the Inter-System Channel (ISC) channels. For example, the InfiniBand physical layer offers the potential to encapsulate ISC data traffic and significantly increase bandwidth between servers (extended distances could be accommodated by WDM or channel extenders). Note that any synchronous remote copy technology will increase the I/O response time, because it will take longer to complete a writing operation with synchronous remote copy than without it (this effect can be offset to some degree by using other approaches, such as parallel access to storage volumes). The tradeoff for longer response times is that no data will be lost or corrupted if there is a single point of failure in the optical network. PPRC makes it possible to maintain synchronous copies of data at distances up to 103 km; however, these distances can only be reached using either DWDM with optical amplifiers or by using some other form of channel extender technology. The performance and response time of PPRC links depends on many factors, including the number of volumes of storage being accessed, the number of logical subsystems across which the data is spread, the speed of the processors in the storage control

units and processors, and the intensity of the concurrent application workload. In general, the performance of DASD and processors has increased significantly over the past decade, to the point where storage control units and processors developed within the past two years have their response time limited mainly by the distance and the available bandwidth. Many typical workloads perform several read operations for each write operation; in this case, the effect of PPRC on response time is not expected to be significant at common access densities.

Similar considerations will apply to any distributed synchronous architecture such as parallel sysplex. In some cases, such as disaster recovery applications where large amounts of data must be remotely backed up to a redundant storage facility, an asynchronous approach is practical. This eliminates the need for Sysplex Timers and trades off continuous real-time data backup for intermittent backup; if the backup interval is sufficiently small, then the impact can be minimized. One example of this approach is the eXtended Remote Copy (XRC) protocols supported by FICON channels on a z series server. This approach interconnects servers and DASD between a primary and a backup location, and periodically initiates a remote copy of data from the primary to the secondary DASD. This approach requires fewer fiber-optic links, and because it does not use a Sysplex Timer the distances can be extended to 100 km or more. The tradeoff with data integrity must be assessed on a case-by-case basis; some users prefer to implement XRC as a first step toward a complete GDPS solution.

The use of a parallel computing architecture over extended distances is a particularly good match with fiber-optic technology. Channel extension is well known in other computer applications, such as storage area networks; today, mainframes are commonly connected to remote storage devices housed tens of kilometers away. This approach, first adopted in the early 1990s, fundamentally changed the way most people planned their computer centers and the amount of data they could safely process; it also led many industry pundits to declare “the death of distance.” Of course, unlike relatively low-bandwidth telephone signals, performance of many data communication protocols begins to suffer with increasing latency (the time delay incurred to complete transfer of data from storage to the processor). While it is easy to place a long-distance phone call from New York to San Francisco (about 42 milliseconds round trip latency in a straight line, longer for a more realistic route), it is impossible to run a synchronous computer architecture over such distances. Further compounding the problem, many data communication protocols were never designed to work efficiently over long distances. They required the computer to send overhead messages to perform functions such as initializing the communication path, verifying it was secure, and confirming error-free transmission for every byte of data. This meant that perhaps a half dozen control messages had to pass back and forth between the computer and storage unit for every block of data, while the computer processor sat idle. The performance of any duplex data link begins to fall off when the time required for the optical signal

to make one round trip equals the time required to transmit all the data in the transceiver memory buffer. Beyond this point, the attached processors and storage need to wait for the arrival of data in transit on the link, and this latency reduces the overall system performance and the effective data rate.

As an example, consider a typical fiber-optic link with a latency of about 10 microseconds per kilometer round trip. A mainframe available in 1995 capable of executing 500 million instructions per second (MIPS) needs to wait for not only the data to arrive, but also for 6 or more handshakes of the overhead protocols to make the round trip from the computer to the storage devices. The computer could be wasting 100 MIPS of work, or 20% of its maximum capacity, while it waited for data to be retrieved from a remote location 20 km away. Although there are other contributing factors, such as the software applications and workload, this problem generally becomes worse as computers get faster, because more and more processor cycles are wasted waiting for the data. As this became a serious problem, various efforts were made to design lower latency communication links. For example, new protocols were introduced, which required fewer handshakes to transfer data efficiently, and the raw bandwidth of the fiber-optic links was increased from ESCON rates (about 17 Mbytes/s) to nearly 100 Mbyte/s for FICON links. But for very large distributed applications, the latency of signals in the optical fiber remains a fundamental limitation; DASD read and write times, which are significantly longer, will also show a more pronounced effect at extended distances.

While the performance of any large-scale computer system is highly application dependent, we can infer some of the effects caused by extended distances. For the case of I/O requests to DASD on an ESCON link, assume that at the primary site a typical storage read or write operation takes 3 ms. The latency of an intersite fiber-optic link is about 10 microseconds/km round trip; this must be multiplied by the intersite distance and the number of acknowledgments required by the data link protocol to determine the impact of intersite distance on performance. If we assume a conservative datacom protocol (such as ESCON) that requires six acknowledgments per operation, then at a distance of 40 km the additional delay is (10 microseconds/km/round trip) (40 km) (six round trips) = 2.4 ms. The time required for a DASD read operation from site B to DASD in site A is then 3 + 2.4 = 5.4 ms. Similarly, a data-mirroring application might require a write operation to the DASD in site A that would then be remote copied to DASD in site B. This operation would take 3 ms for the local write, 2.4 ms latency, and 3 ms for the remote write, or 8.4 ms total. If the data must first be requested from site B before this operation can begin, this adds another 2.4 ms for a total of 10.8 ms. In a similar fashion, performance of all ESCON and HiPerLinks will degrade with distance. There is no general formula to predict this impact; it must be evaluated for each software application and datacom protocol individually.

17.5.1. Time Synchronization in Distributed Computing

Actually, many different scales of time measurement are of interest to computer science and communication systems. Historically, one of the most important applications for highly accurate time synchronization has been precise navigation and satellite tracking, which must be referenced to the Earth's rotation. The timescale developed for such applications is known as Universal Time 1 (UT1). UT1 is computed using astronomical data from observatories around the world; it does not advance at a fixed rate, but speeds up and slows down with the Earth's rate of rotation. While UT1 is measured in terms of the rotation of the Earth with respect to distant stars, it is defined in terms of the length of the mean solar day. This makes it more consistent with civil, or solar, time. Until 1967, the second was defined on the basis of UT1; subsequently the second has been redefined in terms of atomic transitions of cesium-133.[1] At the same time, the need for an accurate time-of-day measure was recognized; this led to the adoption of two basic scales of time: (1), International Atomic Time (TAI), which is based solely on an atomic reference and provides an accurate time base that is increasing at a constant rate with no discontinuities; and (2), Coordinated Universal Time (UTC), which is derived from TAI and is adjusted to keep reasonably close to UT1. UTC is the official replacement for (and generally equivalent to) the better known Greenwich Mean Time (GMT).

Perhaps the most famous computer problem related to timekeeping was the much publicized "Year 2000" problem, but there are other requirements that are less well known. Since January 1, 1972, an occasional correction of exactly one second called a leap second has been inserted into the UTC timescale. It kept UTC time within ± 0.9 seconds of UT1 at all times. These leap seconds have always been positive (although in theory they can be positive or negative) and are coordinated under international agreement by the Bureau International des Poids et Mesures (BIPM) in Paris, France. This adjustment occurs at the end of a UTC month, which is normally on June 30 or December 31. The last minute of a corrected month can, therefore, have either a positive adjustment to 61 seconds or a reduction to 59 seconds. As of January 1, 2006, 23 positive leap seconds have been introduced into UTC. Thus, any timekeeping function used to synchronize computer systems must account for leap seconds and other effects.[2]

---

[1]Specifically, the second is defined by the international metric system as 9,192,631,770 periods of the radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium-133 atom. In 1967, this definition was already 1000 times more accurate than what could be achieved by astronomical methods; today, it is even more accurate.

[2]The effect of a leap second is the introduction of an irregularity into the UTC timescale, so exact interval measurements are not possible using UTC, unless the leap seconds are included in the calculations. After every positive leap second, the difference between TAI and UTC increases by one second.

The Time-of-Day (TOD) clock was first introduced as part of the IBM System/370™ architecture to provide a high-resolution measure of real time. The cycle of the clock is approximately 143 years and wraps on September 18, 2042. In July 1999, the extended TOD clock facility was announced, which extended the TOD clock by 40 bits. This 104-bit value, along with 8-zero bits on the left and a 16-bit programmable field on the right, can be stored by program instructions. With proper support from the operating system, the value of the TOD clock is directly available to application programs and can be used to provide unique timestamps across a Sysplex. Conceptually, the TOD clock is incremented so that a one is added into bit position 51 every microsecond. (In practice, TOD-clock implementations may not provide a full 104-bit counter, but maintain an equivalent stepping rate by incrementing a different bit at such a frequency that the rate of advancing the clock is equivalent.) The stepping rate (The rate at which the bit positions change) for selected TOD clock bit positions is such that a carry out of bit 32 of the TOD clock occurs every 220 microseconds (1.048576 seconds). This interval is sometimes called a mega-microsecond. The use of a binary counter for the time of day, such as the TOD clock, requires the specification of a time origin, or epoch (the time at which the TOD clock value would have been all zeros). The z/Architecture®, ESA/390, and System/370 architectures established the epoch for the TOD clock as January 1, 1900, 0 A.M. GMT.

In the IBM System z/architecture, programs can establish time of day and unambiguously determine the ordering of serialized events, such as updates to a database. The architecture requires that the TOD clock resolution be sufficient to ensure that every value stored by the operating system commands is unique; consecutive instructions that may be executed on different processors or servers must always produce increasing values. Thus the time stamps can be used to reconstruct, recover, or, in many different ways, assure the ordering of serialized updates to shared data.

In a Parallel Sysplex or GDPS, time consistency is maintained across multisystem processes executing on different servers in the same sysplex. This is accomplished through a Sysplex Timer (IBM model 9037), which provides an external master clock (the external time reference or ETR) that can serve as the primary time reference. Synchronization between multiple, redundant Sysplex Timers is maintained through a control link oscillator (CLO) channel. In 2006, IBM withdrew the Sysplex Timer from marketing (although service and support will continue for some time). The replacement method for time synchronization between servers is called the server time protocol (STP), available beginning with the IBM models z9 EC, z9 BC, z990, and z890 servers running z/OS v 1.7 and higher [xx]. This approach further enhanced server time synchronization by enabling scaling over longer distances (up to at least 100 km or 62 miles) and integrated the time distribution function with existing intersystem channel (ISC) peer

mode links. STP can coexist with legacy Sysplex Timer networks and facilitates migration from ETR/CLO links to STP links.

STP is a message-based protocol in which timekeeping information is passed over coupling links between servers, including ISC-3 peer mode links over extended distances and integrated cluster bus (ICB-3 or ICB-4) links within a server. It is recommended that each server be configured with at least two redundant STP communication links to other servers. There is no architectural limit to the maximum number of links that can be defined; instead, this limit is based on the number of coupling links supported by each server in the configuration (the number of links that can be installed varies by server type). Similarly, the maximum number of attached servers supported by any STP-configured server in a CTN is equal to the maximum number of coupling links supported by the servers in the configuration. Not considering redundancy recommendations, this is just the maximum number of combined ISC-3, ICB-3, and ICB-4 links. For initial STP supported systems in 2007, up to 64 combined coupling links are supported; this number may increase in the future. This is an enhancement over the Sysplex Timer, which could only attach to 24 servers and coupling facilities (in high-availability applications, the Sysplex Timer Expanded Availability (EA) configuration is installed, whereby each server and coupling facility attaches to two Sysplex Timers).

STP allows the use of dial-out time services via modem (such as the Automated Computer Time Service (ACTS) or an international equivalent) so that time can be set to an international standard such as UTC or adjusted for leap seconds, local time zones, Daylight Savings Time, and other effects. The CST can be initialized to within +/–100 ms of an external standard; the application must periodically re-dial out (either manually or automatically) to maintain this accuracy.

With the introduction of STP, it became possible to interconnect multiple servers in a hierarchy of time synchronization, leading to several new concepts in timing network design. A coordinated timing network (CTN) contains a collection of servers that are time synchronized to a value called coordinated server time (CST). Thus, CST represents the time for the entire network of servers. All servers in a CTN maintain an identical set of time-control parameters that are used to coordinate the TOD clocks. A CTN can be configured with either all servers running STP (an STP-only CTN) or with the coexistence of servers and coupling facilities using both ETR and STP (mixed CTN). The Sysplex Timer provides the timekeeping information in a Mixed CTN.

The Sysplex Timer distribution network is a star topology, with the Sysplex Timer at the center and time signals emanating to all attached servers. By contrast, STP distributes timing information in hierarchical layers, or strata. The top layer (Stratum 1) distributes time messages to the layer immediately below it (Stratum 2), which in turn distributes time messages to Stratum 3. More layers are conceivably possible, but the current STP implementation is limited to three layers. There

is no way to assign a particular server as a stratum 1, 2, or 3 server. The Stratum 1 level is determined indirectly in one of several ways. In a Mixed CTN, any STP-configured server synchronized to the Sysplex Timer is a Stratum 1 server. Thus, a Mixed CTN is allowed to have multiple Stratum 1 servers. An STP-only CTN must have only one Stratum 1 server; using the server management console, a server must be assigned as the preferred Stratum-1 server or Preferred Time Server. This server should have connectivity to all servers that are destined to be the Stratum 2 servers, either through ISC-3 links in Peer mode, ICB-3, links or ICB-4 links. Typically, a Stratum 2 server is also designated as a backup time server, which takes over in case the Stratum 1 server fails; it has connectivity to the preferred time server, as well as to all other Stratum 2 servers connected to the preferred time server. Thus, determining the number of required STP links is not as straightforward as in an ETR timing network.

Time coordination is also required in other applications besides the Sysplex and Parallel Sysplex configurations, for example, the asynchronous remote copy technology known as z/OS global mirror (previously called extended remote copy (XRC)). In this example, an application I/O operation from a primary or production site is considered to be completed when the data update to the primary storage is completed. Subsequently, a software component called the system data mover (SDM) asynchronously offloads data from the primary storage subsystem's cache and updates the secondary disk volumes at a remote site used for disaster recovery. Data consistency across all primary and secondary volumes spread across any number of storage subsystems is essential for providing data integrity and

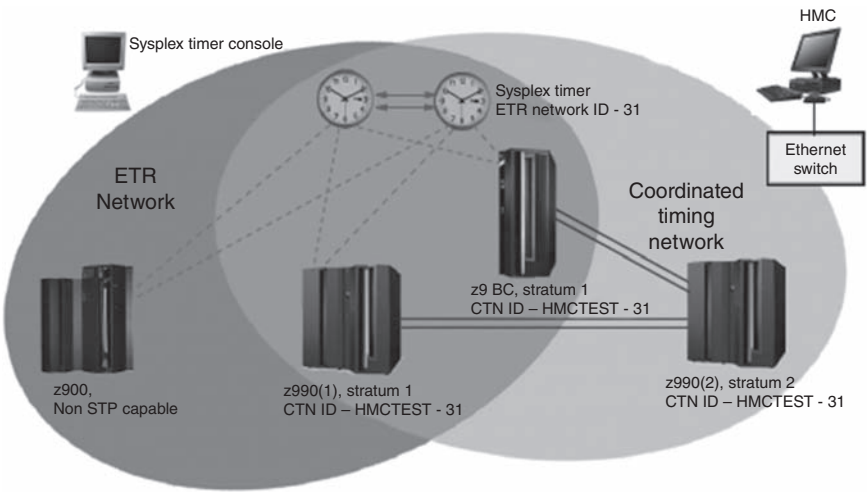


Figure 17.3 Mixed CTN with Stratum 1 and 2 servers [9].

the ability to do a normal database restart in the event of a disaster. Data consistency in this environment is provided by a data structure called the consistency group (CG) whose processing is performed by the SDM. The CG contains records that have their order of update preserved across multiple logical control units within a storage subsystem and across multiple storage subsystems. CG processing is possible only because each update on the primary disk subsystem has been time-stamped. If multiple systems on different servers are updating the data, time coordination using either Sysplex Timer or STP links is required across the different servers in each site. For a server that is not part of a Parallel Sysplex but has to be in the same CTN, additional coupling links must be configured as special "timing-only" links.

17.6. OPTICAL INTERCONNECTS

Figure 17.4 shows the logical hierarchy of interconnects used in server systems. Each individual server typically includes a subset of these buses, although some high-end systems may include them all. Optical interconnects are already frequent in MAN/WAN links and LAN links. This chapter focuses on optical interconnects for cluster links, though much of this discussion will apply to memory bus and SMP (symmetric multiprocessor) bus links.

Table 17.2 contains a list of companies that supply interconnects for various forms of switched fiber. Other companies that supply high-bandwidth (10 gb) bit parallel optical connectors are IntexyS, Corona Optical Systems, OFS, JDSU, and

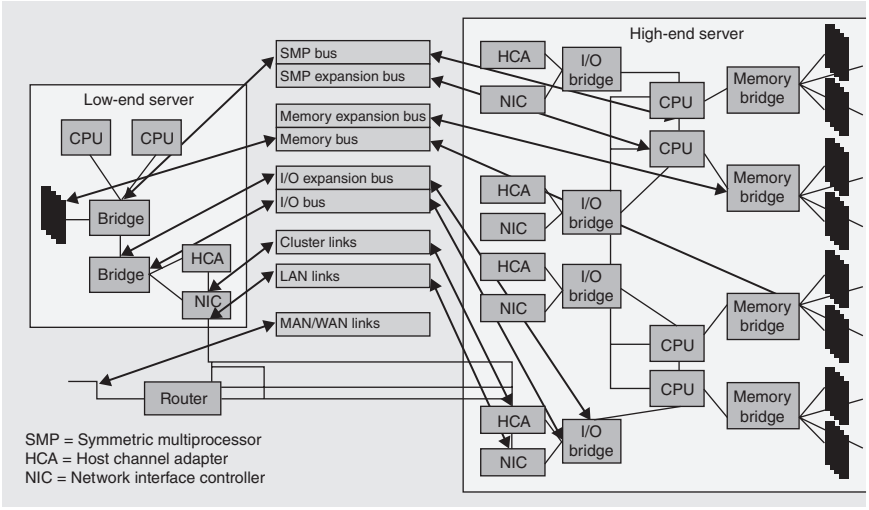


Figure 17.4 Server logical interconnection hierarchy.

Agilent. An important new standard for high-speed optical interconnects is the highest density quad small form-factor pluggable (QSFP) optical module. QSFP connectors are currently supplied by Zarlink, Molex, Emcore, ReflexPhotonics, Tyco Electronics, and Helix Semiconductors. For a more detailed description of QSFP see chapter 3.

17.7 OPTICALLY INTERCONNECTED PARALLEL SUPERCOMPUTERS

Latency is not only a problem for processor to storage interconnections, but also a fundamental limit in the internal design of very large computer systems. Today, many supercomputers are being designed to solve so-called Grand Challenge problems, such as advanced genetics research, modeling global weather patterns or financial portfolio risks, studying astronomical data such as models of the Big Bang and black holes, design of aircraft and spacecraft, or controlling air traffic on a global scale. This class of high-risk/high-reward problems is also known as Deep Computing. A common approach to building very powerful processors is to take a large number of smaller processors and interconnect them in parallel. In some cases, a computational problem can be subdivided into many smaller parts, which are then distributed to the individual processors; the results are then recombined as they are completed to form the final answer. This is one form of asynchronous processing, and many problems fall into this classification; one of the best examples is SETI@home, free software that can be downloaded over the Internet to any home personal computer. Part of the former NASA program, SETI (search for extra-terrestrial intelligence) uses spare processing cycles when a computer is idle to analyze extraterrestrial signals from the Arecibo Radio Telescope, searching for signs of intelligent life. There are currently over 1.6 million SETI@home subscribers in 224 countries, averaging 10 teraflops (10 trillion floating point operations performed per second) and having contributed the equivalent of over 165,000 years of computer time to the project. Taken together, this is arguably the world's largest distributed supercomputer, interconnected mostly with optical fiber via the Internet backbone. Several other problems are being addressed using this model, such as identifying cures for cancer and mapping the human genome.

More conventional approaches rely on large numbers of processors interconnected within a single package. In this case, optical interconnects offer bandwidth and scalability advantages, as well as immunity from electromagnetic noise, which can be a problem on high-speed copper interconnects. For these reasons, fiber-optic links or ribbons are being considered as a next-generation interconnect technology for many parallel computer architectures, such as the PowerParallel and NUMA-Q designs. The use of optical backplanes and related technologies are also being studied for other aspects of computer design (see Chapter 26). To

minimize latency, it is desirable to locate processors as close together as possible, but this is sometimes not possible due to considerations such as the physical size of the packages needed for power and cooling. Reliability of individual computer components is also a factor in how large we can scale parallel processor architectures. As an example, consider the first electronic calculator built at the University of Pennsylvania in 1946, ENIAC (Electronic Numeric Integrator and Computer), which was limited by the reliability of its 18,000 vacuum tubes. The machine could not scale beyond filling a room about 10 by 13 meters, because tubes would blow out faster than people could run from one end of the machine to the other replacing them. Although the reliability of individual components has improved considerably, modern-day supercomputers still require some level of modularity that comes with an associated size and cost penalty.

A well-known example of Deep Computing is the famous chess computer, Deep Blue, that defeated grand master Gary Kasparov in May 1997. As a more practical example, the world's largest supercomputer is currently owned and operated by the U.S. Department of Energy, to simulate the effects of nuclear explosions (such testing having been banned by international treaty). This problem requires a parallel computer about 50 times faster than Deep Blue (although it uses basically the same internal architecture). To accomplish this requires a machine capable of 12 teraflops, a level that computer scientists once thought impossible to reach. Computers with this level of performance have been developed gradually over the years, as part of the Accelerated Strategic Computing Initiative (ASCI) roadmap; but the current generation, called ASCI White, has more than tripled the previous world record for computing power. This single supercomputer consists of hundreds of equipment cabinets housing a total of 8192 processors, interconnected with a mix of copper and optical fiber cables through two layers of switching fabric. Since the cabinets cannot be pressed flat against each other, the total footprint of this machine covers 922 square meters, the equivalent of two basketball courts. This single computer weighs 106 tons (as much as 17 full-size elephants) and had to be shipped to Lawrence Livermore National Labs in California on 28 tractor trailers. It is feasible today to put the two farthest cabinets closer together than about 43 meters, and this latency limits the performance of the parallel computer system. Furthermore, ASCI White requires over 75 terabytes of storage (enough to hold all 17 million books in the Library of Congress), which may also need to be backed up remotely for disaster recovery; so, the effects of latency on the processor-to-storage connections are also critically important.

17.7 OPTICAL INTERCONNECT FUTURES

Current Parallel Sysplex systems have been benchmarked at over 2.5 billion instructions per second, and are likely to continue to significantly increase in

performance each year. The ASCI program has also set aggressive goals for future optically interconnected supercomputers. However, even these are not the most ambitious parallel computers being designed for future applications.

IBM fellow Monty Denneau has led the program to construct a mammoth computer nicknamed "Blue Gene," which will be dedicated to unlocking the secrets of protein folding. Without going into the details of this biotechnology problem, we note that it could lead to innumerable benefits, including a range of designer drugs, whole new branches of pharmacology, and gene therapy treatments that could revolutionize health care, not to mention lending fundamental insights into how the human body works. This is a massive computational problem, and Blue Gene is being designed for the task. When completed, it will be 500 times more powerful than ASCI White, a 12.3-petaflop machine—well over a quadrillion (10^{15}) operations per second, 40 times faster than today's top 40 supercomputers *combined*. The design point proposes 32 microprocessors on a chip, 64 chips on a circuit board, 8 boards in a 6-foot-high tower, and 64 interconnected towers for a total of over 1 million processors. Because of improvements in packaging technology, Blue Gene will occupy somewhat less space than required by simply extrapolating the size of its predecessors: about 11×24 meters (about the size of a tennis court), with a worst-case diagonal distance of about 26 meters. However, the fast processors proposed for this design can magnify the effect of even this much latency to the point where Blue Gene will be wasting about 1.6 billion operations in the time required for a diagonal interconnect using conventional optical fiber. Further more, a machine of this scale is expected to have around 10 terabytes of storage requirements, easily enough to fill another tennis court and give a processor-to-storage latency double that of the processor-to-processor latency. Because of the highly complex nature of the protein folding problem, a typical simulation on Blue Gene could take years to complete and even then may yield just one piece of the answer to a complex protein folding problem.

IBM has recently announced plans to deliver another massively parallel supercomputer to the U.S. government within the next three years, sometimes referred to as "Roadrunner" or "PERCs." This will be the most ambitious supercomputer yet attempted; some of the key planned attributes are shown in Table 17.4. Given the bandwidth and latency requirements, as well as the sheer size of the computer and distance between the furthest separated processors, optical interconnect is expected to play an important role in the realization of this system. The use of parallel optical interconnect may also allow for reduced thermal dissipation and electromagnetic noise emissions.

While designs such as this have yet to be realized, they illustrate the increasing interest in parallel computer architectures as an economical means to achieving higher performance. Both serial and parallel optical links are expected to play an increasing role in this area, serving as both processor-to-processor and processor-to-storage interconnects.

Table 17.4
Historical Evolution of Supercomputer Requirements.

	1995 SP2	2000 ASCI White	2005 ASCI Purple	~2010 PERCS
Max. CPUs	512 POWER2 0.066 GHz	8,192 Power3 0.375 GHz	10,240 Power5 ~2 GHz	524,288 Power7 ~4 GHz
Interconnect BW per link	(0.04+0.04) Gb/s "SP2 Switch"	(0.5+0.5) Gb/s "Colony"	(20+20) Gb/s "Federation"	(120+120) Gb/s "PERCS Network"
Max possible system size	~16 racks	~100 racks	~150 racks	192 racks (incl. 1st-level storage)
BW/rack	(0.2+0.2) T'bits/s 16-switch rack	~(2+2) Terabits/s 16-switch rack	~(10+10) Terabits/s 16-switch rack	~(320+320) Terabits/s (switch+node racks)

ACKNOWLEDGMENTS

The terms System/390, S/390, OS/390, ES/9000, MVS, G5, Parallel Enterprise Server, 9037 Sysplex Timer, Enterprise Systems Connection, ESCON, Fibre Connection, FICON, Parallel Sysplex, HiPerLinks, Fiber Transport Services, FTS, Fiber quick connect, 9729 Optical Wavelength Division Multiplexer, Geographically Dipersed Parallel Sysplex, and GDPS are trademarks of IBM Corporation.

REFERENCES

1. DeCusatis, C., D. Stigliani, W. Mostowy, M. Lewis, D. Petersen, and N. Dhondy. 1999, September–November. Fiber optic interconnects for the IBM S/390 parallel enterprise server. *IBM Journal of Research and Development* 43, no. 5/6: 807–828; also see *IBM Journal of Research & Development*, 36, no. 4 (1992) special issue on IBM System/390: Architecture and Design.
2. DeCusatis, C. 2001, January–February. Dense wavelength division multiplexing for parallel sysplex and metropolitan/storage area networks. *Optical Networks*, pp. 69–80.
3. DeCusatis, C., and P. Das. 2000, January 31–February 1. Subrate multiplexing using time and code division multiple access in dense wavelength division multiplexing networks. *Proc. SPIE Workshop on Optical Networks*, Dallas Texas, pp. 3–11.
4. *Coupling Facility Channel I/O Interface Physical Layer*. Pa.: 1994. (IBM document number SA23–0395). Mechanicsburg, IBM Corporation.
5. Bona, G.I., et al. 1998, December. Wavelength Division Multiplexed Add/Drop Ring Technology in Corporate Backbone Networks. *Optical Engineering*, special issue on Optical Data Communication.
6. *9729 Operators Manual*. 1996, Pa.: IBM Corporation. (IBM document number GA27-4172). Mechanicsburg.
7. IBM Corporation 9729 Optical Wavelength Division Multiplexer. 1996, June. *Photonics Spectra* special issue, the 1996 Photonics Circle of Excellence Awards, vol. 30.

8. DeCusatis, C., D. Petersen, E. Hall, and F. Janniello. 1998, December. Geographically distributed parallel sysplex architecture using optical wavelength division multiplexing. *Optical Engineering*, special issue on Optical Data Communication.
9. Thiébaut, D. 1995. Parallel Programming in C for the Transputer. http://www.cs.smith.edu/~thiebaut/transputer/descript.html?clkd=iwm.
10. Injey, F., N. Dhondy, G. Hutchinson, D. Jorna, G. Kozakos, and I. Neville. 2006, December *Server Time Protocol Planning Guide*, IBM Redbook SG24–7280, 185 pp. , available from www.ibm.com/redbooks.

Relevant Web Sites:

http://www.npac.syr.edu/copywrite/pcw/node1.html is a parallel computing textbook.
http://www.gapcon.com/info.html is a list of the top 500 super computers (all of the current computers referenced were taken from this list).
http://compilers.iecc.com/comparch/article/93-07-068 is a timeline for the history of parallel computing.

This page intentionally left blank

Case Study Parallel Optics for Supercomputer Clustering

Courtesy of IBM Corporation and Avago Corporation

Application: Design one of the world's largest supercomputers capable of solving so-called Grand Challenge problems, including decoding the human genome, protein folding, and nuclear device simulations.

Description: Many emerging applications in "Deep Computing" require scaling computer performance to levels previously thought unattainable (exceeding hundreds of teraflops). In particular, many nations are dealing with the stewardship of their aging stockpile of nuclear devices; due to international test ban treaties, it is no longer possible to detonate such devices either above or below ground, even for test purposes. The need for massive computationally intensive solutions to model these and other problems has driven the interest in recordsetting supercomputer performance. Several years ago, IBM was asked to construct what was at the time the largest supercomputer in the world, based on a nonuniform memory architecture (NUMA) interconnecting thousands of RS/6000 RISC-processor-based compute nodes through a switch fabric (the design called for 8192 processors and 75 terabytes of storage, interconnected through a two layer switch fabric). While previous supercomputing clusters, such as the Earth Simulator in Japan, had relied on lower cost copper interconnects (the large number of links meant that low cost per link was a key design factor), this system introduced the requirement for cascaded switch fabrics and intraswitch links with a greater bandwidth-distance product than was available through copper cables. Low latency was essential for high performance, meaning that serialize/deserialize operations were to be avoided. Instead, 12-channel parallel optical interconnects were designed into the system. Two bidirectional links were implemented on a single adapter blade, using commercially available optical transmitter and

receiver arrays (Agilent/Avago Corporation). This provided 48 GigaBytes per second data throughput (full duplex over ribbons of 24 optical fibers) up to distances of over 100 meters. By contrast, available high-bandwidth copper cables were only capable of about 10 meters; a breakthrough interconnect technology was required in order to construct this computer system. Furthermore, the optical cables provided considerable weight and bulk reduction, which is important considering the sheer scale of this computer. The system completed in June 2000 covered 922 square meters, the equivalent of two basketball courts, and weighed over 106 tons, of which nearly 10 tons was taken up by the copper cables used to interconnect local compute nodes to switches. Reliability of the optical links, in particular the VCSEL lasers, and strain relief of the fiber cables were critical design issues. The final system achieved 12.3 teraflops, more than three times faster than contemporary systems and the first to exceed double-digit teraflop performance.

18

Manufacturing Environmental Laws, Directives, and Challenges

John H. Quick
IBM Corporation

18.1. INTRODUCTION

Manufacturing of optoelectronics technology and components used in fiber-optic data communication products and systems have changed significantly over the course of the last 10 years, and future changes will without doubt continue. In more recent years, this rapid change in fiber-optic technology and manufacturing processes has been motivated to a large extent by the emergence of new worldwide environmental governmental laws and legislative directives. All of these environmental initiatives propose to limit or eliminate heavy metals and other environmental pollutants used in the manufacture of various types of electronic and electric equipment that have been linked to lasting environmental impacts and human health effects.

One of the more significant legislative initiatives, adopted on January 23, 2003, was Directive 2002/95/EC of the European Parliament and of the Council of European Union (EU) on restricting the use of certain hazardous substances in electrical and electronic equipment. [1] The EU directive, nicknamed the Restriction of Hazardous Substances (RoHS), is one of the first to attempt to restrict the use of certain dangerous substances commonly used in electronic and electrical equipment. The RoHS directive is closely linked with the EU Waste Electrical and Electronic Equipment (WEEE) Directive [2, 3], also adopted the same year, which sets collection, recycling, and recovery targets for all types of electrical and electronic goods. The WEEE Directive was adopted in response to the increasing volume of hazardous e-waste being discarded in municipal landfills.

There is no precise definition for e-waste, but it is widely recognized that e-waste includes computer equipment and electronic products used in the data

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

communication (datacom) industry that are broken or not repairable, obsolete, or no longer wanted. Many of the components manufactured for use within electronic equipment contain toxic or hazardous materials that are not biodegradable and that can create serious long-term health risks in the manufacturing of the components and to the natural environment. When equipment is incinerated or discarded haphazardly without pretreatment in municipal and private landfills, these hazardous toxins can be released into the air, water, or land. In 1991 Switzerland was the first country to ban the disposal of e-waste in public lands and landfills in an effort to protect the regions water sources. Recently, the European Union community, states within the United States, and other industrialized countries have enacted similar legislation that require sellers and manufactures of datacom and other related electronic equipment to receive back, reuse, recycle, and otherwise dispose of products using an environmentally responsible process.

Legislative directives and laws enacted now, or working toward approval, have added another level of complexity throughout the components and equipment entire life cycle. These new "green" environmental requirements have and will continue to fundamentally change how companies manage products and conduct business worldwide. In addition, EU directive 2005/32/EC [4], otherwise known as a Directive on Every Energy Using Product (EuP), seeks to create a framework for the integration of different environmental aspects (such as energy efficiency, water consumption, or noise emissions) into the product design of so-called energy-using products to encourage designers and manufacturers to produce products with their environmental impacts in mind throughout the entire product life cycle. EuP, when fully phased in, will require manufacturers to calculate the energy used to produce, transport, sell, use, and dispose of its products. "The [EuP] directive provides for the setting of requirements which the energy-using products covered by implementing measures must fulfill in order for them to be placed on the market." The EuP directive became effective on July 6, 2007. The three referenced directives summarized above are now in effect along with other similar but different laws in force or pending worldwide.

This chapter examines some of the current and pending legislative initiatives, together with the impacts, challenges, and risks that designers, manufactures, and companies will need to consider throughout the product's entire life cycle to produce products that are environmentally friendly. In today's global economy, wide-ranging regulatory measures such as those already mentioned will have a profound impact on a company's operations and the ability to design, manufacture, market, and service information technology equipment worldwide. Companies that ignore these regulations will face stiff monetary penalties for noncompliance, loss of revenue and market share and damage to both client relationships and brand reputation. This section seeks to provide the reader a basic overview of environmental regulation already enacted and does not constitute legal advice. The actual directives, laws, standards, and regulations published in

their original language should always be reviewed and used to ensure product compliance.

18.2. WORLDWIDE ENVIRONMENTAL DIRECTIVES, LAWS AND REGULATIONS

In recent years, countries and regions around the world have been progressively more active in legislating more environmentally friendly and energy-efficient products that are easier to manufacture, recycle, and reuse. These environmental regulations require more open disclosures about the product and their effects on the environment. In this section a comparison (not all inclusive) of five such legislative mandates is summarized, and some of the basics are shown in Table 18.1. Unfortunately, there is no single harmonized standard that companies can use to design and produce products that will meet all worldwide

Table 18.1

Environmental Requirements Comparison Summary.

Parameter	EU	China	California	Japan	Korea
Scope	10 product categories, exclusions	11 product categories	1 product Category	7 product categories	10 products
The Restricted Substances	Lead Cadmium Mercury Hexavalent-Chromium PBB PBDE	Lead Cadmium Mercury Hexavalent-Chromium PBB PBDE	Lead Cadmium Mercury* Hexavalent-Chromium	Lead Cadmium Mercury Hexavalent-Chromium PBB PBDE	Lead Cadmium Mercury Hexavalent-Chromium PBB PBDE
Restriction or Disclosure	Restriction	Disclosure only	Restriction	Disclosure only & Labeling	Disclosure only
Maximum Concentration Values	0.1% for all except cadmium at 0.01%	0.1% for all except cadmium at 0.01%	0.1% for all except cadmium at 0.01%	0.1% for all except cadmium at 0.01%	0.1% for all except cadmium at 0.01%
Level at which restriction is applied	Homogeneous	Homogeneous	Homogeneous	Homogeneous	Homogeneous
Exemptions	Allowed	All EIPs—none. Will be specified in the catalog of listed products	Follows EU	Follows EU	Expected to follow EU

regulations. Although the EU directive has gotten the most notice, the United States and most other developed nations are implementing similar restrictions. The European Union in particular was one of the first to adopt stringent environmental directives, laws, and regulations, including the RoHS Directive banning the use of certain harmful substances; the WEEE directive governing the recovery of waste electronics; and the EuP directive relating to eco-design of products. The EuP legislation is likely to impact datacom designers and products even more than the RoHS and WEEE in that it requires companies to demonstrate they both practice and document eco-design when introducing their products. Companies must do so before product can be sold in Europe.

Today the European Union community consists of Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and United Kingdom. The directive extends to the European Economic Area (EEA), which includes Iceland, Liechtenstein and Norway. Other worldwide countries referenced in Table 18.1 have implemented European-style environmental regulations but with a number of significant differences and additions.

The People's Republic of China promulgated the Management Methods for Controlling Pollution by Electronic Information Products. These methods were developed "to control and reduce pollution to the environment caused after disposal of electronic information products, promote production and sale of low-pollution electronic information products, and safeguard the environment and human health."

"The Law of the People's Republic of China for the Promotion of Clean Production and the Law of the People's Republic of China on the Prevention and Control of Environmental Pollution by Solid Wastes" and associated laws were also enacted [5] Article 1,6,7,8,9,10. Within the industry, this management method is just normally called China RoHS. The method was promulgated on February 28, 2006 and became effective on March 1, 2007. While there are some shared aims between the EU RoHS requirements and those in China, there are also significant differences between them. This China law, like other recent worldwide legislative requirements, was promulgated without the needed guidance necessary for the electronics industry to actually implement it. These deficiencies are normally clarified by the lawmaking agency by publishing an annex or other secondary documents used for compliance guidance. There are key differences between the China and EU RoHS requirements, and these will be examined in the following sections, given that both the EU directive and China RoSH have the largest impact on the design and manufacturing of data communication equipment.

In December 2006, the California Department of Toxic Substances Control (DTSC) adopted emergency regulations to include the RoHS provisions for products sold in California, as established in SB 20 and SB 50. [11] The California

RoHS law, as it is known, consists of two major elements—recycling and restricted substances—and it became effective on January 1, 2007. The California RoHS law only restricts four of the six EU restricted substances in the covered products, which are the heavy metals, lead, mercury, cadmium, and hexavalent chromium. California does not restrict polybrominated biphenyls (PBBs) and polybrominated diphenyl ethers (PBDEs), which are manmade chemicals used as flame retardants mixed into some plastics and other electronic materials within the covered products. All covered electronic devices manufactured after January 1, 2007 are subject to California's RoHS regulations, except for exemptions found in California laws and in the EU RoHS directive or annex. In essence, products that are covered by the California regulations and that are prohibited for sale under the EU directives and exemptions cannot be sold in California. To date, the EU Parliament and Commission have amended the RoHS directive seven times and have enacted 27 exemptions; another 100 are still under consideration. This means that starting on January 1, 2008 American manufacturers that do not sell into the EU will still have to meet the California RoHS "covered electronic device" regulations in their home markets. The California RoHS law, which is found in section 25214.10 of the Health and Safety Code, applies only to a "covered electronic device," which Public Resources Code section 42463 [11] defines as "a video display device containing a screen greater than four inches, measured diagonally."

The covered products are found Appendix X of Chapter 11 of the California Code of Regulations, title 22, as follows:

1. Cathode ray tube containing devices (CRT devices)
2. Cathode ray tubes (CRTs)
3. Computer monitors containing cathode ray tubes
4. Laptop computers with liquid crystal display (LCD)
5. LCD containing desktop
6. Televisions containing cathode ray tubes
7. Televisions containing liquid crystal display (LCD) screens
8. Plasma televisions

California's Integrated Waste Management Board estimates that there are more than 6 million obsolete computer monitors and televisions stockpiled in homes and offices. Electronic devices that do not fall into any of the listed categories are not subject to California's RoHS law. However, California is expected to expand in scope to cover the same products found in the EU directives. California SB 20 and SB 50 include both WEEE and RoHS provisions. The requirements for the recycling and disposal of covered devices became effective on January 1, 2005. Since enactment, clients have paid a recycling fee at the time of purchase on the covered electronic devices. The recycling fee funds e-waste recovery payments to authorized collectors and e-waste recycling payments. At the end of the product's useful life, the client returns the covered e-waste product

to a convenient collection location for recycling. As mentioned earlier, regulations have added another level of complexity throughout the equipment's entire life cycle, especially if the same electronic product such as a monitor is sold in both the commercial datacom and consumer electronic market.

California also mandates that manufacturers of covered electronic products notify the California Integrated Waste Management Board (CIWMB) when a device is subject to the recycling fee. The producer must also provide client information on how to recycle the products and file annual reports with the Board specifying the number of covered devices sold in the state, the total amount of hazardous substances contained in the devices, the company's reduction in use of hazardous materials from the year before, their increase in use of recyclable materials from the year before, and their efforts to design more environmentally friendly products. The United States has no federal legislation parallel to the RoHS directive, though manufacturers are very active now in implementing RoHS compliant technology.

Today many U.S. states and even cities have separate laws in effect or pending for each RoHS defined substance, which presents companies yet another product design and regulation coordination and compliance challenge. For example, a New York City (NYC) bill signed into law on December 29, 2005 mandates that no new covered electronic device purchased or leased by any New York City agency shall contain any six prohibited EU hazardous substances in any amount exceeding that controlled by the director through rulemaking. In developing such rules, the agency director must consider the European Union directive and any subsequent material additions. "Covered electronic device" includes display products.

Japan is similarly establishing environmental legislation in the form of the Law for the Promotion of Effective Utilization of Resources [13].

> The aim of [the] Law for the Promotion of Effective Utilization of Resources is to promote integrated initiatives for the 3Rs (reduce, reuse, recycle) that are necessary for the formation of a sustainable society based on the 3Rs. In particular, it uses cabinet orders to designate the industries and product categories where businesses are required to undertake 3R initiatives, and stipulates by ministerial ordinances the details of voluntary actions that they should take. Ten industries and 69 product categories have been designated, and actions stipulated include 3R policies at the product manufacturing stage, 3R consideration at the design stage, product identification to facilitate separate waste collection, and the creation of voluntary collection and recycling systems by manufacturers, among other topics. [13]

The law was promulgated on June 2000, with enforcement beginning on April 2001.

In November 2005, the Japanese Industrial Standards Committee of the Ministry of Economy, Trade and Industry (METI) issued JIS C 0950:2005, also known as J-MOSS, the Japanese Industrial Standard for Marking the presence Of the Specific chemical Substances for electrical and electronic equipment. This so-called Japanese RoHS standards requirement mandates that manufacturers provide marking and Material Content Declarations for certain categories of electronic products offered for sale in Japan after July 1, 2006. The Japan RoHS standard has EU RoHS like requirements (the same six restricted substances and the same maximum concentration levels) but uses a voluntary approach for compliance rather than a legislative mandate. The J-MOSS requirements apply to personal computers (including LCD and CRT displays) and many other commercial and consumer target product groups. A key element in this standard requires mandatory product labels using the J-MOSS mark label and reporting, effective July 1, 2006. The standard requires that manufacturers and import sellers of target products manage the six specified RoSH substances if contained in the target product. When the product's content exceeds the values set in the standards, the manufacturers must display the "content mark," which is a two-hand clasping "R" symbol on the product and packaging (Fig. 18.1) and the substance information must be disclosed in catalogs and instruction manuals, as well as the Internet.

The content mark indicates that the specific chemical substance should be managed in the Supply chain for proper recycling. The green content mark, which is a two-hand clasping "G" symbol (Fig. 18.2), is optional for electrical and electronic equipment and can be used when the content rates of all the specified chemicals are equal to or less than the stand value of content rates; part of the content chemicals are exempt from content marking; and the other content rate of the other specified chemicals is equal to or less than the standard content value.

Japan RoHS content mark labels should be made from a durable material with a permanent adhesive to ensure that it will last the life of the product. The purpose of the marking is to properly sort out and manage the products throughout the

Figure 18.1 J-MOSS Orange "Content Mark."

Figure 18.2 J-MOSS Green "Content Mark."

reuse/recycle stage. Japan has also issued similar regulations, for instance, the Law Concerning the Protection of the Ozone Layer through the Control of Specified Substances and Other Measures (Law No. 53 of May 20, 1988) (Japan), which focuses on eliminating various classes of ozone gasses.

South Korea adopted the Act for Resource Recycling of Electrical and Electronic Equipment and Vehicles "for the promotion of recycling of, design for the environment of, and restriction of hazardous substances in electrical and electronic equipment and vehicles and appropriate treatment of their waste. This Act is to contribute to the preservation of the environment and healthy development of national economy through the establishment of a resource reduction, reuse and recycling system for efficient use of resources" [16]. The Korean RoHS and recycling legislation compliance data is set for January 1, 2008. Again, there are some similarities between the Korea RoHS and the China RoHS, particularly with regards to providing clear compliance guidance. Today, the official legislation is short on detailed requirements within the document. The Korea Ministry of Environment has indicated that the RoHS restrictions will be consistent with the European Union and China directives.

One divergence between the China RoHS and the Korea RoHS is that Korea is not requiring product labeling. However, manufacturers will be responsible for collecting and managing the material composition data that shows their compliance. The minister of environment and the minister of commerce, industry and energy will determine and publish methods for analyzing hazardous substances. The analysis methods have not yet been published. Companies will also have to register and declare that products comply with the Korea Act.

Korea RoHS's recycling provision differs from the EU's WEEE directive. The Act will require posting of the required documents to an Operations Management Information System in lieu of paper recordkeeping. The Korea Ministry of Environment (MoE) will establish an electronic management system to allow manufacturers to post data electronically. The Korea Act also provides for public officials to inspect business places, facilities, equipment, and documents at any time to verify compliance with the Act. Notice is given seven days prior to the inspection. The intent of the Korea Act is to start first with recycling electronic

and electrical products and automobiles and perhaps expand over time. Companies must be constantly aware of new and changing environmental, recycling, and EuP legislation.

18.3. RESTRICTION OF HAZARDOUS SUBSTANCES (RoHS)

RoHS is a European Union (EU) RoHS directive that aims to restrict certain dangerous substances commonly used in electrical, electronic components, and electronic equipment. As stated, "*The purpose of this Directive is to approximate the laws of the Member States on the restrictions of the use of hazardous substances in electrical and electronic equipment and to contribute to the protection of human health and the environmentally sound recovery and disposal of waste electrical and electronic equipment" [1]*. It is important to understand that the European Union RoHS directive is not a law but rather a legislative act that requires member states to accomplish a particular result without dictating how they must achieve the directive's objective. As with many European directives, enforcement is the responsibility of each individual country in the EU, which decides the preferred methods of enforcement and the penalties that will be levied against the manufacturer for noncompliance of the electronic equipment sold on the EU market after the July 1, 2006 deadline. This section reviews the main RoHS requirements; readers are encouraged to read the official referenced documents and their appendixes.

Remember! The European Union's ROHS directive is NOT the same as China's RoHS, Japan's RoHS, Korea's RoHS, or any other RoHS. As mentioned earlier, there is no regulatory harmonization.

The restriction on the use of certain hazardous substances (RoHS) directive (often referred to as the lead-free directive) restricts six substances. The six substances and their maximum concentration are shown in Table 18.2.

The maximum concentration limit is calculated by weight at the raw "homogeneous material" level, which is a unit (not the finished product or a component)

Table 18.2

E.U. RoHS Restricted Materials and Maximum Concentration Levels.

RoHS Restricted Materials	Maximum Concentration Limits*
Lead (Pb)	0.1% by weight or 1000 ppm
Mercury (Hg)	0.1% by weight or 1000 ppm
Cadmium (Cd)	0.01% by weight or 100 ppm
Hexavalent Chromium (CrVI)	0.1% by weight or 1000 ppm
Polybrominated Biphenyls (PBB)	0.1% by weight or 1000 ppm
Polybrominated Diphenyl Ethers (PBDE)	0.1% by weight or 1000 ppm

of any single substance that cannot be mechanically disjointed into different materials. The term *homogeneous* will not be found in any of the EU directives or in the Guide to the Implementation of Directives Based on the New Approach and the Global Approach (commonly referred to as the 'blue book') [17]. However, the United Kingdom (UK) government and Commission suggest the term *homogeneous* to be understood as "of uniform composition throughout." So examples of "homogeneous materials" would be individual types of plastics, ceramics, glass, metals, alloys, paper, board, resins, and coatings. The UK commission goes on to suggest that "mechanically disjointed" means that the materials can, in principle, be separated by mechanical actions such as unscrewing, cutting, crushing, grinding, and abrasive processes. Take a fiber-optic transceiver, for example. It consists of optical subassemblies, chips on PCB board, integrated circuit drivers (die), metal ferrules, and plated terminal pins, all contained within a metal or plastic housing. The transceiver as a whole is not homogeneous since it can clearly be separated using the methods described above. In essence, the legislation applies to the lowest common denominator of an item of uniform composition. Any single identifiable one of the six RoSH substances in Table 18.2 must not be present in the homogenous material above the maximum concentration values, unless covered by an exemption. In addition, the substance mercury must not be intentionally added to any component. If any material exceeds the maximum limit, then the entire component or product wherein the substance is used would fail the EU directive and could not be "put on the market." The "put on the market" expression comes from Article 4.1 of the RoHS Directive, which states: "*Member States shall ensure that, from 1 July 2006, new electrical and electronic equipment put on the market does not contain lead, mercury, cadmium, hexavalent chromium, polybrominated biphenyls (PBB) or polybrominated diphenyl ethers (PBDE). National measures restricting or prohibiting the use of these substances in electrical and electronic equipment which were adopted in line with Community legislation before the adoption of this Directive may be maintained until 1 July 2006.*" [1].

The RoHS-style environmental legislation written in Europe, China, California, Japan, and Korea, is slightly different for each region. The differences are visible in critical areas such as exceptions, reporting, and proof of compliance. The matrix of varying environmental compliance rules seen in Table 18.1 will get more complicated as new rules such as the European Union's REACH [x] laws restricting additional chemicals emerge. Manufacturing of optoelectronics technology and components used in fiber-optic data communication products and systems are most affected by the EU RoHS and China RoSH regulations. As new legislation emerges, manufacturers will have to determine what laws carry the strictest rules and comply with those laws. Table 18.3 compares the EU and China RoHS requirements and provides a good view of the complexities involved in meeting product environmental compliance.

Table 18.3
A Comparison of EU RoHS and China RoHS.

Subject Area	EU RoHS	China RoHS
Legislation adopted	February 13, 2003	February 28, 2006
Effective Date (in force)	July 1, 2006	March 1, 2007
Scope	Ten broad categories of finished products. Individual product types are not specified and legislation leaves interpretation to producer [1]	All Electronic Information Products (EIP). Extensive list published which includes many products exempt and not covered by EU RoHS such as medical equipment, measurement instruments, some production equipment, batteries, and most types of components
Main requirements	Six RoHS substances are restricted and must not be present in homogeneous materials, at above the maximum concentration values, unless covered by an exemption Table 18.2	Two levels of requirements: All EIPs must be marked to indicate whether any of the six substances are present. Products that will be specified in a catalog—substance restrictions will be specified, and these may be some or all of the six EU-RoHS substances and possibly others
Restricted substances	Lead, cadmium, mercury, hexavalent chromium, PBB, and PBDE	None—disclosure, reporting, and labeling only for EU RoHS substance above limits
Marking and disclosure	None—Related WEEE directive requires use of the crossed wheelie bin symbol to indicate to users that the product should be recycled at end of life	Four Requirements: (1) Disclose hazardous materials and locations (2) Environment-friendly use period Label (3) Packaging materials mark (4) Date of manufacture
Sources of details of legislation	Published EC and member state guidance and some Commission Decisions [1]	Chinese Standards to be published by Chinese government and some Q & A from Ministry of Information Industry [5]
Maximum concentration values	In-scope products must contain less than: 0.1% for all except Cd which is 0.01%. All are by weight in homogeneous materials (unless covered by exemptions) Table 18.2	Marking with a table and the orange logo if concentrations of Pb, Hg, Cr(6), PBB, or PBDE are >0.1% or >0.01% of Cd by weight in homogeneous materials, except for metal coatings where RoHS substances must not be intentionally added and parts of 4 mm_3 or less regarded as single homogeneous materials
Exemptions	29 granted, more than 70 under consideration	None—All EIPs are specified in catalog for listed products
Testing/Certification and approach to compliance	Self-declaration, third-party testing not required	Self-declaration for marking of all IEPs. Testing by authorized laboratories in China of catalog listed products

Table 18.3 (*continued*)

Subject Area	EU RoHS	China RoHS
Legislation adopted	February 13, 2003	February 28, 2006
Effective Date (in force)	July 1, 2006	March 1, 2007
Packaging	Not included as covered by the Packaging Directive: European Parliament and Council Directive 94/62/EC of 20 December 1994 on packaging and packaging waste	Must be nontoxic and recyclable and marked to show materials' content
Batteries	Not included, covered by EU batteries and accumulators directive	Included within EIPs catalog
Nonelectrical products	Excluded if the finished product sold to user does not depend on electricity for its main function	Included if listed as EIPs. Includes CDs and DVDs
Products used for military and national security use only	Excluded from EU scope	Excluded from China scope
"Put onto the market"	When product is made available for first time sale within EU and transferred to distribution	Applies to products produced on or after March 1, 2007 and must be marked from date forward

Both the EU and China have legal regulations of comparable intent to recycle and control hazardous substances in electronic and electrical equipment by controlling the concentration values. From this point forward both regulations differ, as shown in the comparison chart.

One of the principal differences between the EU and China RoHS is the China Marking for Control of Pollution Caused by Electronic Information Products (SJ/T 11364-2006) requirements. This standard describes labeling requirements in detail. Although China RoHS does not require the removal of hazardous substances, the law requires the manufacturers to label the product and provide a table in the user's guide disclosing the location of any hazardous substance above the maximum concentration values (MCVs). The next step is to calculate the Environmentally Friendly Use Period (EFUP) value.

The EFUP value is defined in the ACPEIP (Administration on the Control of Pollution caused by Electronic Information Products) [5] in Article 3 as "*The term during which toxic and hazardous substances or elements contained in*

Table 18.4
Hazardous Substance Disclosure Table.

有毒有害物质或元素名称及含量标识样式						
	有毒有害物质或元素					
部件名称	铅 (Pb)	贡 (Hg)	镉 (Cd)	六价铬 (Cr^{6+})	多溴联苯 (PBB)	多溴二苯醚 (PBDE)
机架 chasis	0	0	0	X	0	0
处理器模块 processor modules	X	0	0	0	0	0
逻辑模块 logic modules	X	0	0	0	0	0
电缆组合件 cable assemblies	X	0	0	0	0	X
监视器 monitor	0	0	0	X	0	0

electronic information products will not leak out or mutate, thus eliminating the possibility of serious environmental pollution resulting from the use by users of electronic information products or serious harm to their persons and properties resulting from such use". The EFUP Draft Standard of August 20062 [18] five methods for calculating the EFUP, split into two categories.

Technical based EFUP

1. The Practical Method
2. Experimental Method

Theoretical based EFUP

3. Technical Life Method
4. Safe Use Period Method
5. Comparison Method

It is stated in the EFUP Draft General rule August 20062 that if the technical based EFUP is known then this should be used. The equipment producer must determine the EFRUP using one of these methods. For details of methods that can be used, please see the China RoHS Guidance Notes available from *RoHS-International* or a recent translation of the EFUP Guidance available from *Design Chain Associates*. The equipment manufacturer must detail the method used, and any assumptions for determining the EFUP in the user's manual. Detailing the calculation method used is not a legal requirement but it is considered a good business practice considering the variability of the methods. Once the EFUP value is determined for the product, the legislation requires the product be labeled and dated with one of the two Pollution Control Marks.

example of the Pollution Control Mark I Logo

Pollution Control Mark I Logo (also indicates recyclables) is used when there are no RoHS substances present at concentrations greater than the maximum concentration levels (same six as EU RoHS except Deca-DBE).

example of the Pollution Control Mark II Logo

Pollution Control Mark II Logo is used when there are hazardous substances present at concentrations greater than the maximum concentration levels. The number within the mark is Environment Friendly Use Period (in years). Further, the legislation requires that the label needs to be located in a location visible to the user and can be molded, painted, stuck or printed on the product. The date of manufacture must also be printed on the product.

EU WEEE “wheelie bin” label

The EU WEEE directive requires that all products be marked with the “wheelie bin” symbol to indicate that they may not be discarded for curbside pick up.

As described in this section, equipment and component manufactures face significant challenges to design manufacture and ship environmental compliant products worldwide. Without worldwide RoSH and recycling harmonization legislation, Datacom equipment manufactures will continue struggle with costly processes to comply individual legislation for the reason that each nations scope is different; the requirements are different; some have included exemptions others do not; and yet other require labels, marks, and disclosure if their products contain hazardous substances. In addition, the concept of “Put on the market” is different, the penalties for noncompliance are different and the responsibilities dictated by the law are different.

Components and equipment suppliers will also need to be responsive to OEM clients that may have environmental requirements that are more stringent than those required by current governmental legislation. For example, International Business Machines (IBM) requires suppliers to conform to “IBM Engineering Specification (ES 46G3772) which establishes the baseline environmental requirements for supplier deliverables to IBM. This requirement along with other IBM specifications, contracts and procurement documents contain additional environmental requirements for suppliers. ES 46G3772 [19] contains restrictions on materials in products and on certain chemicals used in manufacturing. It also

requires suppliers to disclose information about the content of certain materials in their products. In addition, the specification includes requirements for batteries, marking of plastic parts, and other product labeling requirements. [20]

18.4 ENVIRONMENT REQUIREMENT COMPLIANCE

How do you know your product is compliant? Will your documentation withstand examination? Will your product prove to actually be compliant if it is taken part and tested? What documents have you got to offer to the authorities if they challenge your product declaration? These are just a few of the questions that Datcom equipment and component manufactures will have to consider. Compliance will require that manufactures and their component suppliers understand the material composition of their products. This includes bulk materials, individual components, sub assemblies and finished products. Equipment and component manufactures must also have and retain detail technical documentation to support their declarations in support of their "due diligence," Manufactures will be expected to provide this documentation upon request of regulators. Most of the legislative mandates require or strongly suggest that all "reasonable steps and due diligence" have taken to avoid any regulatory offense. This also implies that some amount of testing may be needed to ensure product compliance. Manufactures will have to carefully assess and select parts that have the highest probability of containing restricted hazardous substance Not every country intends to have a due diligence compliance declaration defense. Many will make the offense one of strict liability. If the IT equipment contains banned hazardous substances beyond allowable levels, the producer will be guilty. Other countries have adopted a mix form of strict liability, with the penalty varying depending whether the manufacturer is considered to be negligent. There is no single solution to demonstrate "Due Diligence". However, manufactures will require suppliers to provide conformation of compliance documents or to provide material content declarations similar to IPC 1752. The "IPC 1752 for Material Declarations" [21] is the standard for the exchange of materials declaration data focused on printed circuit board assemblies. A group of Original Equipment Manufactures (OEMs), Electronic Manufacturing Services (EMS) providers, component manufacturers, circuit board manufacturers, materials suppliers, information technology solution providers, and the National Institute of Standards and Technology developed the IPC 1752 standard. Since each Datacom producer will want some appropriate information, the standard has established 6 classes of disclosure. There is no definitive guidance on what exactly will be considered to be all reasonable steps, but manufactures should consider strict supply chain managements methods, compliance testing, third party evaluations, a data base for materials or products or other third party certification like ECO Labeling ("Green Seal"). Opposite to the EU RoHS approach to material content self-certification, the China RoHS law

will require a product to be tested before it is allowed entry into China, and only testing by Chinese certified labs will be accepted by the Chinese authorities. The China legislation covers all categories of optical communications and attachment equipment.

The EEE industry estimates that the average cost of IT systems to support and demonstrate compliance with environmental initiatives at $2 million to $3 million per company, with deployment time of a year or more.

Another RoHS accepted "screening" method practiced in the industry is X–ray Fluorescence testing which is an analytical process widely used for quantitative materials testing. These instruments use safe x-ray sources to fluoresce characteristic x-rays from materials. By analyzing the energy of the fluoresced x-rays the unit can determine what elements are present in the material being analyzed and approximate the element's concentration. These units can probably tell if a product is in gross violation of RoHS but XRF should not be used for definitive results; since for example there is no speciation of Chromium (Cr+6) and Bromine (PBB/PBDE). Material testing down to the homogeneous materials in every single part you use to build your product may be required, but in reality is not realistic. However, China does require "proof" that products are compliant. The Chinese authorities don't have to prove to the producer that they are not. Today, the best practice compliance process is to collect information from the supply chain for each component, verify collected information and fill in compliance gaps and then store, audit, and update information from the supply chain. Manufactures should also conduct a risk assessment for each supplier to determine the accuracy of the provided information. Suppliers considered "High Risk" should be asked to provide independent third party test results. Third party importers, wholesalers, distributors and retailers have to accept responsibility for shipped product. If product labeling and documentation is inadequate, wrong or unsupported, they risk the same sanctions as the producer.

18.5 ENVIRONMENT BUSINESS IMPACTS

The implications of RoHS, reuse and recycling legislation on the datacom industry is enormous. There are many business and process issues that electronic and datacom equipment suppliers have to do to guarantee RoHS compliance and to limit potential legal responsibility. Compliance will require traditional program management techniques, internal and supplier communication, education, participation and cooperation among all of the functions needed to produce a product. EEE manufactures need to establish roadmap and compliance strategies and processes to manage supply chain implications and detailed product analysis. Manufactures will need to "know the law" and conduct compliance "gap analysis" while monitoring new regulatory developments and requirements.

Global warming, depleting resources, the impact of hazardous substances, and waste disposal have all become high profile subjects in the last few years and are

starting to have substantial impacts on the fiber optic datacom and electronic component industry.

Environmental regulations have changed how products are designed, manufactured and reclaimed and their reach is expected to expand beyond the current requirements. It is estimated that by 2010 most developed companies will adapt similar governmental laws and additional directives. Other legislations such as the Energy using Products (EuP) and REACH Directives are likely to impact industry even more than the WEEE and RoHS as it requires companies to demonstrate they both practice and document eco-design when launching their products. They must do so before they can sell their products in Europe.

This EuP Directive encourages the eco-design of equipment that uses less energy throughout its lifecycle and to avoid the use of hazardous materials, not only in the products but also in the manufacturing process of raw materials and component parts. The REACH directive deals with the **R**egistration, **E**valuation, **A**uthorisation and Restriction of **Ch**emical substances. [22] This legislation is also expected to affect equipment design and makes it more difficult for designers to justify the use of toxic substances in materials in the product. The new law entered into force on June 1, 2007 Designing for environmental compliance goes beyond checking to make sure a particular component or product meets legislative rules. The entire process involves making sure the component and supporting products comes with its appropriate materials declaration so the manufacturer can prove to governmental entities that govern compliance that they took all appropriate measures to make sure the product was designed to comply.

Throughout this chapter we examined some of the current and pending legislative initiatives; impacts, challenges and risks that designers, manufactures and companies need to consider throughout the products entire life cycle actions to produce products that are environmentally friendly. This is not a one-time effort, but an ongoing set of activities that fiber optics datacom component and equipment producers will be facing for a long time forward.

REFERENCES

1. Official Journal L 037, 13/02/2003 P. 0019–0023, Index 32002L0095, *Directive 2002/95/EC of the European Parliament and of the Council of 27 January 2003 on the restriction of the use of certain hazardous substances in electrical and electronic equipment*, http://ec.europa.eu/environment/waste/weee/legis_en.htm.
2. Official Journal L 037, 13/02/2003 P. 0024–0039, Index 3200210096, *Directive 2002/96/EC of the European Parliament and of the Council of 27 January 2003 on waste electrical and electronic equipment (WEEE)—Joint declaration of the European Parliament, the Council and the Commission relating to Article* 9 http://ec.europa.eu/environment/waste/weee/legis_en.htm.
3. Official Journal L 345, 31/12/2003 P. 0106–0107–0039, Index 32003L01, *Directive 2003/108/EC of the European Parliament and of the Council of 8 December 2003 amending Directive 2002/96/EC on waste electrical and electronic equipment (WEEE)*, http://ec.europa.eu/environment/waste/weee/legis_en.htm.

4. (Official Journal L 191, 22.7.2005, p. 29–58), *Directive 2005/32/EC of the European Parliament and of the Council of 6 July 2005 establishing a framework for the setting of ecodesign requirements for energy-using products and amending Council Directive 92/42/EEC and Directives 96/57/EC and 2000/55/EC of the European Parliament and of the Council*, http://ec.europa.eu/enterprise/eco_design/dir2005–32.htm.
5. People's Republic of China—Management Methods for Controlling Pollution by Electronic Information Products, English: http://www.aeanet.org/governmentaffairs/gabl_ChinaRoHS_FINAL_March2006.asp
6. People's Republic of China—Ministry of Information Industry—Electronic Information Products Classification and Explanation, English: http://www.aeanet.org/governmentaffairs/gabl_HK_Art3_EIPTranslation.asp
7. People's Republic of China SJ/T 11363-2006 Requirements for Concentration Limits for Certain Hazardous Substances in Electronic Information Products, http://www.aeanet.org/governmentaffairs/gajl_MCV_SJT11363_2006ENG.asp
8. People's Republic of China SJ/T 11364-2006 Marking for Control of Pollution caused by Electronic Information Products, http://www.aeanet.org/governmentaffairs/gajl_LABELING_SJT11364_2006ENG.asp
9. People's Republic of China SJ/T 11365-2006 Testing Methods for Toxic and Hazardous Substances in Electronic Information Products (draft version), http://www.aeanet.org/governmentaffairs/gajl_ChinaRoHS_TestingMethods_August2006.asp
10. People's Republic of China GB 18455-2001 Packaging Recycling Mark, http://www.aeanet.org/governmentaffairs/gajl_Packaging_GB18455_2001ENG.asp
11. California Department of Toxic Substance Control, Laws Regulations and Policies, http://www.dtsc.ca.gov/LawsRegsPolicies/
12. The Law Concerning the Examination and Regulation of Manufacture etc. of Chemical Substances (1973 Law No. 117, last Amended July 2002) substances from products, http://www5.cao.go.jp/otodb/english/houseido/hou/lh_04050.html
13. Japan Law, Law for the Promotion of Effective Utilization of Resources, http://www.meti.go.jp/policy/recycle/main/english/law/promotion.html
14. Japan Law, The Law Concerning the Examination and Regulation of Manufacture etc. of Chemical Substances (1973 Law No. 117, last Amended July 2002) substances from products, http://www5.cao.go.jp/otodb/english/houseido/hou/lh_04050.html
15. Guide to the Implementation of Directives Based on New Approach and Global Approach, http://ec.europa.eu/enterprise/newapproach/legislation/guide/index.htm
16. Korea Law, "Act for Resource Recycling of Electrical and Electronic Equipment and Vehicles", April 2007, http://www.kece.eu/data/Korea_RoHS_ELV_April_2007_EcoFrontier.pdf
17. Guide to the Implementation of Directives Based on New Approach and Global Approach, http://ec.europa.eu/enterprise/newapproach/legislation/guide/index.htm
18. AeA translation of the August 29 2006 "General rule of Environmental-Friendly Use Period of Electronic Information Products" http://www.aeanet.org/governmentaffairs/gabl_EPUP_Guidelines_Aug_2006.asp
19. "Engineering Specification 46G3772: Baseline Environmental Requirements for Supplier Deliverables to IBM" http://www.ibm.com/ibm/environment/products/especs.shtml
20. List of IBM Documents Referenced in ES 46G3772 and Information for suppliers and import compliance guidelines, http://www-03.ibm.com/procurement/proweb.nsf/ContentDocsByTitle/United+States~Information+for+suppliers
21. Association Connecting Electronics Industries, IPC 1752 for Materials Declaration, http://members.ip"c.org/committee/drafts/2-18_d_MaterialsDeclarationRequest.asp
22. European Commission REACH information page, http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm

19

ATM, SONET, and GFP[1]

Carl Beckmann
Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire

Rakesh Thapar
Marconi, Warrendale, Pennsylvania

19.1. INTRODUCTION

Early communication networks were driven by telephony and telecommunications requirements. In the beginning, there were analog networks for supporting telephony. In the 1960s it became apparent that delivering analog telephone calls using analog frequency-division multiplexing techniques was prone to noise and did not make as effective use of the available bandwidth on copper wires as could digital techniques. Then, long-haul digital networks were installed for delivering long-distance telephone service. (Local service remained primarily analog for some time.)

For telephony, the basic requirements are to establish a point-to-point connection for a "call" that is typically several minutes in duration. The delay through the network must be small enough not to interfere with the quality of speech and to avoid perceptible echo effects, on the order of milliseconds or less. In North America, voice is digitized to 8 bits precision at 8000 samples per second (to support an analog bandwidth of 300–3000 Hz), for a data rate of 64 kb/s per call. The capacity of a typical copper wire digital trunk link (a T1 line) is 1.5 Mb/s by comparison.

These basic requirements are served well by using reserved connections along all the links from a source to a destination (circuit switching), with time-division multiplexing (TDM) used to share the bandwidth among multiple calls on the individual links. TDM keeps the latency on each link very low while dividing the

[1]Material on GFP has been added by the editors for this edition.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

available bandwidth evenly between calls, with only a small amount of framing overhead. American and international standards bodies have adopted a variety of standard data rates and interoperability specifications, which are summarized in Appendix D.

19.1.1. Data Communications and Packet Switching

Although telecommunications networks can be used to carry data as well, local area data networks can be built much more cheaply and efficiently without resorting to switch-based network architectures. The requirements for data networks come from the ability to transfer files and other packets of information, such as electronic mail messages and terminal data interfaces, consisting of typed keyboard strokes from a user and displayed textual and graphic information back to the user. Other traffic comes from remote procedure calls for distributed computing and operating system information and distributed file system transfers. Compared with telephony, a much more heterogeneous mix of traffic exists on data networks.

For most data traffic, there are no hard real-time constraints on its delivery. File transfers should happen quickly to allow for smooth operation and rapid response time to interactive users; slower response time is merely annoying but does not render the service useless.

In early systems, network bandwidth and delay were dictated by near-real-time requirements of character terminal input/output: A fast typist can type 60 to 100 words per minute. If each word is an average of 6 characters long (including the space), and each character is represented in 8-bit ASCII, then this represents a steady throughput of 80 bits per second. Characters come at an average rate of 10 per second. Terminal equipment was usually connected to the computer (often through an intermediate multiplexer) via a 9600-baud serial data line. Thus, the events of interest (defining the maximal acceptable latency) are on the order of hundreds of milliseconds, and the bandwidth required per connection is on the order of tens of kilobits or less per second. One hundred users could be accommodated with approximately 1 Mb/s worth of usable bandwidth.

Due to the highly heterogeneous and unpredictable nature of traffic on data networks, the use of connection-oriented communications, in which bandwidth is reallocated only every few minutes per channel, is not efficient. Moreover, higher latencies are tolerable. This makes connectionless communications based on discrete data packets, each containing its own addressing and format information, much more attractive. The bandwidth of the network can, in effect, be reallocated on a packet-by-packet basis on demand as the packets arrive at the network. To make this scheme efficient, a larger number of data bits should be put in each packet that is required for the packet “header” information. This incurs extra latency but is tolerable.

19.1.2. Asynchronous Transfer Mode and Synchronous Optical Network Overview

ATM has been proposed as an enabling network technology to support broadband integrated services. It is not a complete, stand-alone networking standard. Rather, ATM defines a common layer of interoperability called the ATM layer, on which various services ranging from telephony and video conferencing to TCP/IP data networking and multimedia can be delivered. The ATM layer defines a common format used for switching and multiplexing bit streams from one end of an ATM network to another.

The ATM layer, in turn, uses the hardware facilities of lower layers to deliver the bits across individual links in a network. A variety of such physical layers have been defined, most of which are based on existing standards in order to maximally leverage existing technologies and installed bases. These relationships are summarized in Fig. 19.1.

One family of ATM physical layers is based on SONET—a synchronous, time-division multiplexing standard based on transmission over optical media (actually, a family of standards at a variety of bit rates). It was designed primarily to support telecommunications and long-haul, broadband services.

19.2. SONET

19.2.1. Historical Perspective

The SONET standards were developed in the mid-1980s to take advantage of low-cost transmission over optical fibers. SONET defines a hierarchy of data rates, formats for framing and multiplexing the payload data, as well as optical

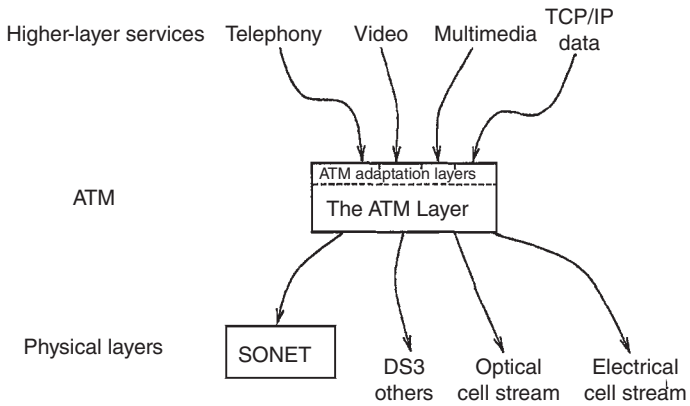


Figure 19.1 ATM and SONET in perspective.

Table 19.1
Basic SONET/SDH Data Rates.

SONET			
Electrical	Optical	SDH	Rate
STS-1	OC-1	—	51.840 Mb/s
STS-3	OC-3	STM-1	155.520 Mb/s
STS-9	OC-9	—	466.560 Mb/s
STS-12	OC-12	STM-4	622.080 Mb/s
STS-18	OC-18	—	933.120 Mb/s
STS-24	OC-24	—	1.244160 Mb/s
STS-36	OC-36	—	1.866240 Mb/s
STS-48	OC-48	STM-16	2.488320 Mb/s

signal specifications (wavelength and dispersion), allowing multivendor interoperability. SONET was originally proposed by Bellcore in 1985 and later standardized by ANSI and the CCITT [synchronous digital hierarchy (SDH) is a compatible set of standards in Europe] [l–3].

SONET is designed to support existing telephone network trunk traffic and also designed with broadband ISDN (BISDN) services in mind. Its TDM basis readily supports fixed-rate services such as telephony. Its synchronous nature is designed to accept traffic at fixed multiples of a basic rate, without requiring variable stuff bits or complex rate adaptation. The SONET data transmission format is based on a 125-μs frame consisting of 810 octets, of which 36 are overhead and 774 are payload data. The basic SONET signal, whose electrical and optical versions are referred to as STS-1 and OC-1, respectively, is thus a 51.84 Mb/s data stream that readily accommodates TDM channels in multiples of 8 kb/s.

SONET defines a hierarchy of signals at multiples of the basic STS-1 rate. The SONET rates currently standardized are shown in Table 19.1. SDH is a compatible European counterpart to SONET. Due to compatibility issues with European switching equipment, the basic SDH rate, called STM-1, is three times the STS-1 rate (i.e., STS-3), or 155.52 Mb/s.

19.2.2. STS Data Rates and Framing

To efficiently support telephony, SONET bit rates rest fundamentally on voice-quality audio sampling rates, that is, 8000 samples per second at 8 bits per sample. The SONET data transmission format is therefore based on a 125-μs frame illustrated in Fig. 19.2. This figure shows the basic STS-1 frame. Higher rates are achieved by byte-interleaving multiple STS-1 frames. The 125-ps frame contains 6480 bit periods, or 810 octets (bytes).

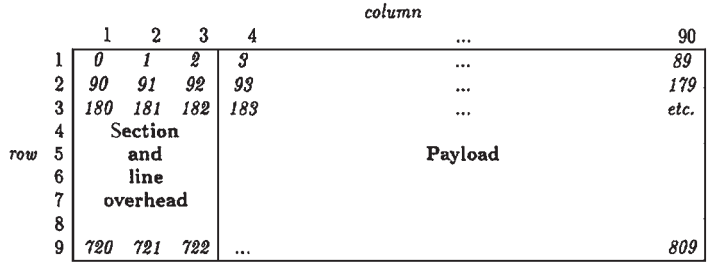


Figure 19.2 SONET framing format.

This can be viewed as a two-dimensional arrangement of nine rows by 90 columns (of bytes) that is scanned row-wise from the upper left.

Thus, a single-voice channel occupies a single octet in each 125-μs frame, and after leaving room for various "overhead" octets (see below), 774 64 kb/s voice channels can be time-division multiplexed into a single STS-1 frame. The bit rate for an *OC-N* link is thus given by

$$\begin{aligned} \text{OC-N bitrate} &= \text{N} \cdot 8000\,\text{Hz} \cdot 90\ \text{columns} \cdot 9\ \text{rows} \cdot 8\ \text{bit/octet} \\ &= 51.84\,\text{N}\,\text{Mb/s}, \end{aligned} \tag{19.1}$$

and the payload capacity (after accounting for four overhead columns per frame) is

$$\text{OC-N capacity} = \text{N} \cdot 8000\,\text{Hz} \cdot 86.9 \cdot 8 = 49.536\,\text{N}\,\text{Mb/s}. \tag{19.2}$$

The first three columns in each frame (i.e., the first three of every 90 octets) are reserved for various overhead bytes. Overhead information is organized into section, line, and path overhead.

SONET can be thought of as following a layered model. At the lowest layer (the physical layer), SONET specifies characteristics of the optical signal, such as maximum dispersion. The lowest level physical link between two pieces of SONET equipment (i.e., an optical fiber pair) is called a section (Fig. 19.3). Multiple sections may be linked together via signal repeaters (regenerators) to form a line. The two ends of a line attach to line termination equipment. At the next level up, a physical line may be used by one or more paths, which are connected on both ends to path termination equipment. It is at the path termination equipment that SONET frames are assembled and disassembled. The layered approach allows the use of equipment for handling functions related to one or the other layer individually, keeping costs down by not requiring all layers to be handled at once.

The first three columns of each frame contain the section and line overhead bytes. The first three rows of this are for the section overhead, and the last six rows are for the line overhead. This is illustrated in Fig. 19.7. The remaining 87

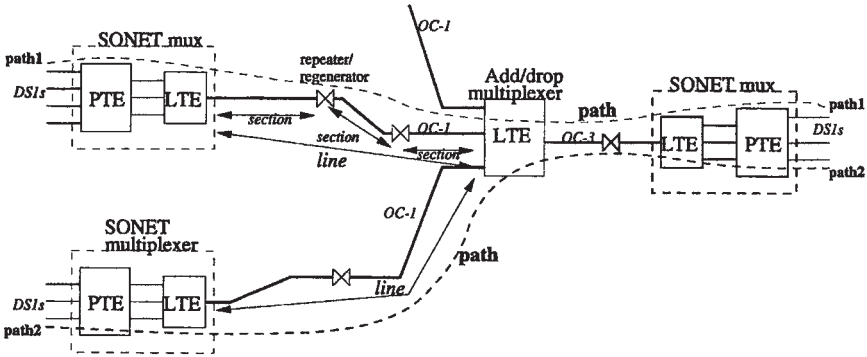


Figure 19.3 SONET sections, lines, and paths.

Table 19.2
SONET Frame Overhead Bytes.

		1	2	3	4			90
section overhead	1	A1	A2	C1				
	2	B1	E1	F1				
	3	D1	D2	D3				
line overhead	4	H1	H2	H3		path overhead		
	5	B2	K1	K2		J1		
	6	D4	D5	D6		B3		
	7	D7	D8	D9		C2		
	8	D10	D11	D12		G1		
	9	Z1	Z2	E2		F2	payload	
next frame						H4		
						Z3		
						Z4		
						Z5		

columns contain the synchronous payload envelope (SPE). The SPE contains the actual payload data as well as a single column of path overhead bytes.

Note that the SPE need not be exactly aligned in the payload frame. In fact, the first byte of the SPE may reside (and usually does) anywhere within the frame; hence, the path overhead is not always in column 4. Overhead octets H1 and H2 form a pointer to the location of the first SPE octet. This feature is useful in connecting two lines whose bit clocks differ slightly, as they do in practice. This allows the SPE to "slip" slightly with respect to the frame. A stuff byte is provided

Table 19.3
SONET Section Overhead Octets.

Symbol	Bits	Name	Description
A1, A2	16	Framing	F628 Hex (1111011000101000 binary); provided in all STS-1 signals within an STS-N signal
C1	8	STS-1 identification	Unique number assigned just prior to interleaving that stays with STS-1 until deinterleaving
B1	8	Section BIP-8	Allocated in each STS-1 for a section error monitoring function
E1	8	Orderwire	Used as a local orderwire channel; reserved for communications between regenerators, hubs, and remote terminal locations
F1	8	Section user channel	This byte is set aside for the network provider's purpose; it is passed from one section level entity to another and is terminated at all section-level equipment
D1–D3	24	Section datacomm	A 192-kb/s channel for alarms, maintenance, control, etc. between section terminating equipment

in H3 to make up the bandwidth deficit in the case in which the signal to transmit is faster than the line clock. This scheme separates the synchronization of data payload frames from the generation of the framing signals, which can be done from a transmitter's local clock.

19.2.2.1. Section Overhead Octets

The first three rows of the first three columns in each frame are used for section-related functions. The functions of these bytes, which include framing, identification, section error monitoring, and auxiliary data channels, are summarized in Table 19.3.

19.2.2.2. Line Overhead Octets

The last six rows in the first three columns of each frame are used for line-related functions, as summarized in Table 19.4.

19.2.2.3. Path Overhead Octets

The first column in the SPE of an STS-1 signal is used for various path-related functions, as summarized in Table 19.5. In an OC-N signal, which carries *N* byte-interleaved STS-1 SPEs, the first column in each STS-1 is used for path-related overhead. By contrast, in a "concatenated" OC-Nc signal, there is only a single column of path overhead, with the remaining 87N-1 columns available for payload data.

Table 19.4
SONET Line Overhead Octets.

Symbol	Bits	Name	Description
H1, H2	16	Pointer	Indicates the offset in bytes between the pointer and the first byte in the STS SPE
H3	8	Pointer action	Stuff byte for downstream frame advancement
B2	8	Line BIP-8	Allocated in each STS-1 for a line error monitoring function; used as a local orderwire channel; reserved for communications between regenerators, hubs, and remote terminal locations
K1, K2	16	APS channel	Allocated for APS signaling between two line-level entities; also carried other management signals
D4–D12	72	Line datacomm	Nine bytes (576 kb/s) allocated for line data communication for alarms, maintenance, control, etc.
Z1, Z2	16	Growth	Further expansion
E2	8	Orderwire	Express orderwire between line entities

Table 19.5
SONET Path Overhead Octets.

Symbol	Bits	Name	Description
J1	8	STS path trace	Used by path-terminating equipment to verify its connection to the source, which continuously sends a fixed 64-byte pattern
B3	8	Path BIP-8	Path error monitoring
C2	8	STS path signal label	Indication of valid construction of SPE
G1	8	Path status	Path-terminating status and performance, back to an originating path
F2	8	Path user channel	For network provider
H4	8	Multiframe	A 192-kb/s channel for alarms, maintenance, control, etc., between section terminating equipment
Z3–Z5	24	Growth	Further expansion

19.2.3. Payload Envelope Pointer

The SPE of a SONET frame need not be perfectly aligned with the framing overhead. Pointer octets HI and H2 are used to locate the SPE within the frame. The lower 10 bits of H1 and H2 are an offset to the beginning of the SPE, that is, the number of octets between H3 and J1, the first octet in the SPE. This feature makes it easier to synchronize multiple signals and multiple pieces of equipment, while allowing each signal source to generate its own framing structure based on a local clock.

The upper 4 bits of H1 and H2 are used to signal changes in the pointer value: A value of 0110 signals an increment or decrement by 1, and a value of 1001 signals some larger change. In the frame in which the pointer is incremented by 1, the lower 10 (H1, H2) bits do not contain the new pointer value but rather the old pointer value, with all the even bits (including the LSB) inverted; on a decrement by 1, the odd bits are inverted. Once the pointer stabilizes, the true new value is used in the lower 10 (H1, H2) bits.

Because the frequency deviation imposed by the standard is small, pointer adjustments will take place infrequently in practice. If an upstream clock is too slow, the downstream equipment will have to periodically increment its pointer and delay outgoing SPEs. When eventually the pointer overflows the maximum value of 809, an entire frame will be skipped. If the upstream clock is too fast, the pointer will have to be decremented periodically. When this happens, the missing byte is put in the H3 octet to compensate. Essentially, the H3 stuff byte provides the extra bandwidth needed for slow-running clocks to keep up with the required data rate.

19.2.4. Multiplexing

Higher speed transmission than STS-1 rates is achieved by byte-interleaving N STS-1 signals to obtain an STS-N signal (which is then converted to an optical OC-N signal). This allows, for example, several STS-1 signals to be multiplexed for transmission over an OC-3 (or higher) link.

Alternately, higher speed channels can be obtained using concatenated STS-1s to achieve a single channel with N times the capacity of an STS-1. In this case, N STS-1 frames are again byte-interleaved to obtain the STS-Nc framing structure. In the STS-Nc frame, there are 3N columns for transport (section and line) overhead, with 87N columns remaining for the payload. However, this payload is multiplexed, switched, and transported through the network as a single entity. Hence, only a single column of path overhead is needed (leaving slightly more bandwidth available for data capacity compared to noncontatenated STS-N).

19.2.5. Virtual Tributaries

In order to directly support services with lower bandwidth requirements than the basic STS-1 payload, several standard “virtual tributary” formats have been defined for SONET. These are summarized in Table 19.6. The VT1.5, for example, allows a DS 1 or T1 signal to be carried end to end on a SONET path without having to remultiplex the *24* DSO (voice) channels contained therein.

Each virtual tributary format is defined as some integral number of columns of the SONET SPE, which includes room for the carried signal as well as any VT-related overhead octets.

Table 19.6
Virtual Tributaries.

Name	Service	Data rate	No. columns
VT1.5	DS1	1.544	3
VT2	CEPT	2.048	4
VT3	DS1C	3.088	6
VT6	DS2	6.176	12

19.2.6. International Interoperability

SONET is compatible with an international set of standards called the SDH. SDH was developed based on SONET, but with the additional goal of providing compatibility between North American and European telecom carriers. Whereas SONET starts with a 51.84 Mb/s signal consisting of nine rows by 90 columns every 125 μs (STS-1), SDH starts with a 9×270 frame every 125 μs, or a 155.52 Mb/s signal.

The basic 155.52 Mb/s SDH signal, called STM-1, is similar and can be made compatible with SONET STS-3. There are some differences in the usage of section and line overhead octets between SONET and SDH. For a more detailed discussion of the differences, the reader is referred to Minoli [3]. See also Table 19.1 for SDH data rates.

19.2.7. Sonet Physical Specifications

Specifications for the transmitter, receiver, and optical signal path characteristics for various SONET line rates are given in Table 19.7 [4].

19.3. ATM

19.3.1. Cell vs. Packet Switching

ATM is designed for high-speed transport of a variety of traffic types. Due to its high-speed nature, it is believed that using fixed-size cells will allow efficient hardware implementations of various multiplexing and routing functions. Unlike LAN environments using Ethernet, fast Ethernet, or fiber distributed data interface (FDDI), and capable of tens to hundreds of megabits of throughput on variable-sized packet traffic, ATM is designed to work into the gigabit per second range [2, 3, 5–7]. Moreover, ATM is designed to be able to support both switched- and packet-oriented applications. Finally, ATM allows the quality of service to be specified within a range of parameters during call setup time.

The use of small, fixed-sized cells has several advantages over larger variable-sized packets as used in Ethernet or FDDI:

Table 19.7

SONET Physical Layer Optical Specifications.

Parameter	Units	OC-1	OC-3	OC-9	OC-12	OC-18	OC-24	OC-36	OC-48
Data rate									
Bit rate	Mb/s	51.84	155.52	466.56	622.08	933.12	1244.16	1866.26	2488.32
Tolerance	ppm			100					
Transmitter									
Type		MLM/LED	MLM/LED	MLM/LED	MLM/LED	MLM	MLM	MLM	MLM
$\lambda_{W\min}$	nm	1260	1260	1260	1260	1260	1260	1260	1265
$\lambda_{W\max}$	nm	1360	1360	1360	1360	1360	1360	1360	1360
$\Delta_{\lambda\max}$	nm	80	40/80	19/45	14.5/35	9.5	7	4.8	4
$P_{T\max}$	dBm	–14	–8	–8	–8	–8	–5	–3	–3
$P_{T\min}$	dBm	–23	–15	–15	–15	–15	–12	–10	–10
$r_{e\min}$	dB	8.2	8.2	8.2	8.2	8.2	8.2	8.2	8.2
Optical path									
System									
$ORI_{\max}$	dB	na	na	na	na	20	20	24	24
$D_{SR\max}$	ps/nm	na	na	31/na	13/na	13	13	13	12
Max sndr. to revr. reflectance	dB	na	na	na	na	–25	–25	–27	–27
Receiver									
$P_{R\max}$	dBm	–14	–8	–8	–8	–8	–5	–3	–3
$P_{R\min}$	dBm	–31	–23	–23	–23	–23	–20	–18	–18
P_0	dBm	1	1	1	1	1	1	1	1

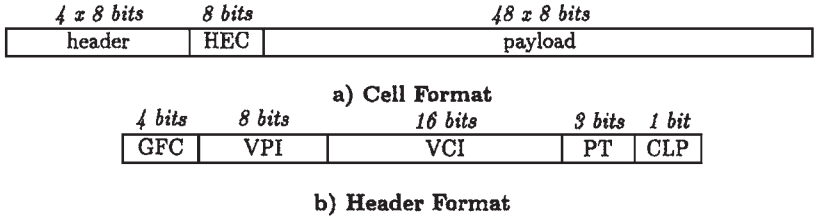


Figure 19.4 ATM cell format.

- Cell boundaries can be easily recognized at high speed in hardware, should loss of framing occur.
- Individual packets cannot monopolize the bandwidth of the channel.
- Cell-handling decisions (e.g., during congestion or for traffic policing of individual connections) can be easily made based solely on the number of cells, without having to examine their headers for packet size information.
- Cell-buffering hardware in switches and other equipment is simplified.
- Circuit-like switching of cells replaces store-and-forward routing of packets, with much lower latency over multihop paths.

The disadvantages are that header information may consume a larger fraction of available bandwidth than for large packets, and that sending very small amounts of information is less efficient than it is for small packets (although both are inefficient).

The structure of ATM cells is shown in Fig. 19.4 [8]. It consists of a 5-byte header followed by 48 bytes of payload data. The header contains the following fields: generic flow control (GFC), virtual path identifier (VPI), virtual channel identifier (VCI), payload type (PT), a cell loss priority bit (CLP), and header error correction (HEC). A brief description of each of these fields is given in Table 19.8.

19.3.2. Cell vs. Circuit Switching

Another key feature of ATM is its ability to transport "constant bit rate" data such as (uncompressed) telephony or video over virtual circuits with guaranteed bandwidth and latency characteristics. In other words, ATM provides a service that mimics a point-to-point, synchronous connection normally provided by a TDM network. Features of ATM that enable this include the following:

- Cell size is kept small, because cell size directly affects latency at the source and destination associated with packing and unpacking a bit stream into cells and, to some extent, affects latency of cell handling at network-switching elements.

Table 19.8
ATM Cell Fields.

Field	Bits	Name	Description
GFC	4	Generic flow control	
VPI	8	Virtual path identifier	Identifies 1 of 256 possible paths out of the current swith or device. Used with VCI to distinguish and locally route different cell streams
VCI	16	Virtual channel identifier	Identifies 1 of 65,536 possible channels in the given path out of the current switch or device. Used with VPI to distinguish and locally route different cell streams
PT	3	Payload type	Differentiates control vs. data cells, etc.
CLP	1	Cell loss priority	Used to mark low-priority cells that may be discarded if network traffic is high
HEC	8	Header error correction	A CRC checksum on the first four header octets, using the generator polynomial $x^8 + x^2 + x + 1$. The resulting code is also XORed with 01010101 to get the HEC bits
Data	48×8	Payload	User data and headers/trailers from higher network layers

- Keeping cell size fixed makes it feasible to allocate link bandwidth to individual connections, and reducing the cell size increases the bandwidth resolution at which this can be done.
- Fixed-size cells make scheduling of periodic or pseudoperiodic traffic at switching elements feasible in principle.

However, the main justification for ATM is its ability to mix synchronous with other types of traffic such as variable bit rate or connectionless and "bursty" data: By using cells, rather than TDM time slots, the channel bandwidth can be reallocated to different "virtual connections" on a cell-by-cell basis instead of requiring a TDM time slot to be allocated (requiring an end-to-end call setup) as short-lived connections come and go or as the bandwidth requirement of a single channel waxes and wanes. This makes efficient statistical multiplexing possible, where a large number of variable bandwidth connections can be supported over a broadband channel with capacity for the sum of the connections' average bandwidth requirement, even though the sum of maximum instantaneous bandwidth requirements may exceed the channel's capacity.

Paradoxically, one of the physical layers used for ATM is the SONET. Here, SONET frames are simply used to transport ATM cells across a SONET path. The payload carried by the ATM cells need not be synchronous, however, because from frame to frame (and from cell to cell), the payload carried by the ATM cells can come from completely different ATM connections. The ATM cells in the

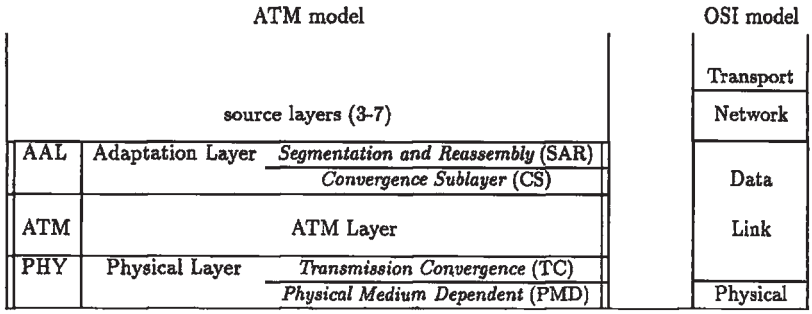


Figure 19.5 ATM reference model.

SONET payload are opaque to the SONET layer. Allocation of the ATM cell bandwidth to CBR, VBR, and connectionless data channels is handled entirely at the ATM layer and above.

19.3.3. ATM Layered Architecture

ATM is based on a layered architecture (Fig. 19.5). The major layers are the physical layer, the ATM layer, and the ATM adaptation layer (AAL). Above the AAL reside the data source layers, corresponding approximately to open systems interconnection (OSI) layers 3–7. The physical layer is further divided into a lower physical medium-dependent sublayer (PMD) and the transmission convergence sublayer (TC). The adaptation layer is also divided into the segmentation and reassembly sublayer (SAR) and the convergence sublayer (CS).

ATM layers do not correspond to the standard seven-layer OS1 model, although an approximate correspondence is shown in Fig. 19.5. In most applications, AAL, ATM, and the TC sublayer of PHY can be thought of as providing the functionality of the OS1 data link layer, that is, the error-free transmission of bits from one end of a link to another. Although this may involve the traversal of several switches (which in turn uses routing information in cells' VPI and VCI fields), the actual network layer function of establishing these routes on call setup is left to higher layers.

The ATM layer is fixed, but a variety of adaptation layers and physical layers have been defined. The services provided by the adaptation layer depend on the traffic type being supported. Traffic types vary in their data rate characteristics (constant data rate versus variable or bursty data traffic), connectionless (datagram) versus connection-orientedness, allowing ATM to support the spectrum of services including voice, video, and computer data services and interactive multimedia. Four basic traffic classes have been defined as shown in Table 19.9, and an adaptation layer has been defined for each (AAL-1-AAL-4). (The adaptation layers

Table 19.9
ATM Service Class.

			ABR	
	CBR (class 1)	VBR (class 2)	Class 3	Class 4
Timing	Synchronous	Synchronous	Asynchronous	Asynchronous
Bit rate	Constant	Variable	Variable	Variable
Connection mode	Connection oriented	Connection oriented	Connection oriented	Connectionless

for class 3 and class 4 available bit rate traffic have been combined into a single layer, AAL-3/4.)

A fifth adaptation layer, AAL-5 (originally called SEAL, the simple and efficient adaptation layer), has also been defined to serve as a convenient application programmer interface (API) for computer applications to build directly on top of ATM services.

19.3.4. ATM Physical Layers

ATM is a switching and multiplexing scheme for BISDN, but it is not necessarily tied to a particular physical layer. Fiber optic as well as electronic physical layers are possible at a variety of data rates [8–11]. At the time of this writing, the ATM Forum Technical Committee has standardized the following physical layers for ATM: 155 and 622 Mb/s fiber-optic layers based on SONET; 100 and 155 Mb/s cell-stream fiber-optic layers; 155 and 25 Mb/s layers for twisted-pair connections; and a DS1 (1.5 Mb/s) layer based on T1. The physical layers standardized for ATM are summarized in Table 19.10 [12].

The ATM user-network interface (UNI) specification [8] includes two kinds of interfaces; public and private. Public ATM service providers and any equipment connecting to public ATM networks must adhere to the public UNI specification, whereas the less stringent private UNI specification is suitable for use in local area networking equipment. The private UNI does not need the operation and maintenance complexity or the link distance provided by the public UNI for telecom lines.

19.3.4.1. SONET/SDH

SONET-based fiber-optic physical layers for ATM have been defined at 155 Mb/s (OC-3) and 622 Mb/s (OC-12) rates. In both of these cases, the PMD sublayer is essentially identical to the corresponding SONET-SDH specification. The TC sublayer makes use of SONET framing by encapsulating ATM cells into the SONET SPE.

Table 19.10
Standardized ATM Physical Layers.

Rate (Mb/s)	Media	Framing	UNI Specification
1.554	Twisted pair	DS1	Public
2.048	Twisted pair, coax	E1	Public
6.312	Coax	J2	Public
25.6	UTP-3	Cell stream, 32 Mbaud	Private
34.368	Coax	E3	Public
44.736	Coax	DS3	Public
51.84	UTP-3	SONET STS-1	Private
100	MMF	Cell stream, 125 Mbaud	Private
155.52	SMF	SONET OC-3c	Public/Private
155.52	UTP-3, coax	SONET STS-3c	Private
155.52	MMF	Cell stream, 194.4 Mbaud	Private
155.52	STP	Cell stream, 194.4 Mbaud	Private
622.08	SMF, MMF	SONET OC-12	Private

The SONET payload envelope presents a bandwidth resource that is used by the TC sublayer to carry ATM cells. However, because the ATM cell size (53 octets) does not evenly divide the size of either the STS-3c or STS-12c payload envelopes, no synchronization between ATM cells and the SONET framing structure is implied (i.e., cells may cross SONET frame boundaries). In the STS-3c UNI, the available capacity for ATM cells is nine rows by 260 columns (the payload envelope minus one column of path overhead), or 149.760 Mb/s. In the 622 Mb/s interface, there are three fixed stuff columns following the path overhead, so the available cell carrying capacity is $9 \times (1044 - 4)/125\,\mu s = 599.04$ Mb/s. In both cases, the available capacity is packed with ATM cells, and any rate decoupling between the ATM and PHY layers is accomplished by inserting empty cells into the stream.

Because of the asynchrony, the TC sublayer is also responsible for cell delineation. This is accomplished via the HEC bits in the cell headers. If cell synchronization is lost, the TC sublayer receiver continuously scans the SONET payload, testing whether each new octet starts a 5-octet ATM header with a valid HEC field. If so, it enters a presynch state, and if several valid cells are detected in a row, the synch state is entered and cell synchronization is assumed. HEC checking continues during normal transmission to verify that cell synchronization is not lost.

As long as cell synchronization is maintained in steady state, the HEC field is also used to correct any single-bit errors found in individual cell headers. The HEC field uses apolynomial code as indicated in Table 19.8 to perform single-bit correction and multiple-bit error detection on the header portion of each cell.

Finally, prior to insertion in the cell stream (and after removal on the receive side), the TC sublayer scrambles the payload portion of ATM cells to avoid any

problems with DC levels or repeated bit patterns in the SONET payload envelope. This uses a self-synchronizing scrambler polynomial described in ITU recommendation 1.432 [8].

19.3.4.2. Cell Stream

Alternately, cells may be sent directly over optical media, without using SONET framing. Several such physical layers have been defined for ATM at 100 and 155.52 Mb/s data rates.

In the cell-stream interfaces, the TC sublayer is responsible for the functions of cell delineation and HEC verification and for the 155 Mb/s UNI, 125-μs clock recovery. In both of these private UNIs, the HEC is used for detection of errored cells only and not for correction because the use of a 4B/5B (or 8B/10B) code means that any line bit errors result in multiple data bit errors. ATM cells are simply discarded from the stream sent to the ATM layer.

The 100 Mb/s TC sublayer interface (also called TAXI) has no framing structure; when no cells are being transmitted, a special 8-bit symbol is continuously sent (not the FDDI "idle line" code).

Although ATM cells are available from the ATM layer, they are transmitted on the line continuously as 54 FDDI symbol pairs each. The 155 Mb/s TC sublayer interface, on the other hand, does have a framing structure consisting of 1 ATM cell used as a physical layer overhead unit (PLOU) followed by 26 cells of data. All 27 cells consist of 53 bytes each, coded as a single 8B/10B symbol as specified by the Fibre Channel standard (see Chapter 20). Unlike the 100 Mb/s interface, cell-rate decoupling is performed by inserting idle cells rather than some idle line symbols. As in the case of SONET-based PHYs, the available bandwidth is packed with a contiguous stream of whole cells. The 155 Mb/s TC sublayer also delivers a 125 microsecond clock across the link, using a special line code (K28.2), which can be inserted anywhere within the symbol stream. This is removed from the symbol stream at the receive end prior to ATM cells.

19.3.4.3. Physical Media Requirements

The optical ATM physical layers include those based on SONET-SDH and direct cell-stream physical layers. Currently, all public UNI specifications for fiber-optic transmission are based on SONET-SDH; hence, they are suitable for long-distance links. The optical specifications can be found in Table 19.7.

The non-SONET cell-stream PHYs have been approved only for the private UNI. These are intended for shorter distance links (up to 2 km) in LANs. The 100-Mb/s layer is based on the physical specifications for FDDI (see Chapter 23).

Table 19.11
Effective Payload Capacity Comparison.

	SONET		ATM		
	OC-3	OC-3c	OC-3c	155 Mb/s	100 Mb/s TAXI
Baud rate	155.52	155.52	155.52	194.4	125
Bit rate	155.52	155.52	155.52	155.52	100
Payload capacity	148.608	149.760	135.6317	135.6317	88.889
Total efficiency	95.56%	96.30%	87.21%	69.77%	71.11%

19.3.4.4. Payload Capacity Comparison

It is instructive to compare the delivered payload capacity of various SONET and ATM formats. Table 19.11 gives the baud rate (line data rate), the bit rate (nominal symbol rate), and delivered payload capacity of raw SONET and several ATM PHY layers in the nominal 100–155 Mb/s range. In the ATM case, the payload capacity listed is the total data payload (no cell headers or HEC) presented to the ATM layer. In the case of SONET, it is the synchronous payload envelope minus any path overhead bytes. In terms of overall efficiency, raw SONET requires only approximately 4% overhead but delivers only synchronous data. ATM incurs an additional 10% of overhead (5 cell header bytes per 48 bytes of data)—the price for the added flexibility of cell versus synchronous TDM switching.

The ATM-cell-stream formats lose 20% in overhead due to the 4B/5B or 8B/10B line symbol coding (compared to 4% for some of the same functionality provided by SONET). Note that the SONET OC-3c and 155 Mb/s cell-stream ATM PHYs have the same effective payload capacity by design, although the forms of overhead are different (SONET framing versus one PLOU per 26 cells).

19.3.5. ATM Layer

As discussed previously, ATM can be implemented atop a variety of physical layers. On the other hand, ATM supports a variety of different services and traffic classes by providing different adaptation layers to higher network levels. The ability to do so efficiently over a shared infrastructure is made possible by a common middle layer, called the ATM layer.

ATM does not use the standard OS1 seven-layer reference model, but the ATM layer performs many of the functions of OS1 level 2, the datalink layer. For example, the ATM layer and AAL-5 together provide datalink layer functionality similar to OS1 layer 2, that is, the ability to transmit error-free frames of size up to 64,000 bytes from a source to a destination ATM entity [5].

The ATM layer is responsible for the switching and multiplexing of ATM cells. Because ATM is based on switched point-to-point links, as opposed to a broadcast medium, ATM functionality is basically connection oriented (although connectionless services are supported through an adaptation layer). Although the ATM layer performs the basic operations that transmit cells along multilink paths from source to destination, the establishment of these paths [i.e., network-layer routing using, for example, an Internet protocol address] is left to higher layers. The ATM layer is designed for simplicity and for ease of high-speed hardware implementation.

19.3.5.1. Virtual Channels and Paths

A virtual channel is a contiguous stream of cells transmitted between two points in an ATM network (e.g., a single user's data stream). To reach the destination from the source, this data stream must traverse a set of ATM switches, going out a particular port on each switch to reach the next switch. This constitutes a virtual path, and many virtual channels may share the same virtual path. Virtual channels can be thought of as being contained inside virtual paths (Fig. 19.6).

Each ATM cell header has 24 bits for identifying the VC and VP that a cell belongs to, the VCI and VPI fields, respectively. This information ultimately is used to route the cell to the correct output port on each switch in its intended path. The generic local routing procedure at each switch is as follows: When a cell enters an input port, its header is examined and the VCI and/or VPI field is extracted. This information is used to index a look-up table, which (i) identifies the output port to send the cell to and (ii) provides new values to be placed into the outgoing cell's VCI and/or VPI fields.

Hence, the VCI and VPI fields are not constant, but rather change on each hop through the network. A virtual path (channel) is a string of VPI (VCI) values

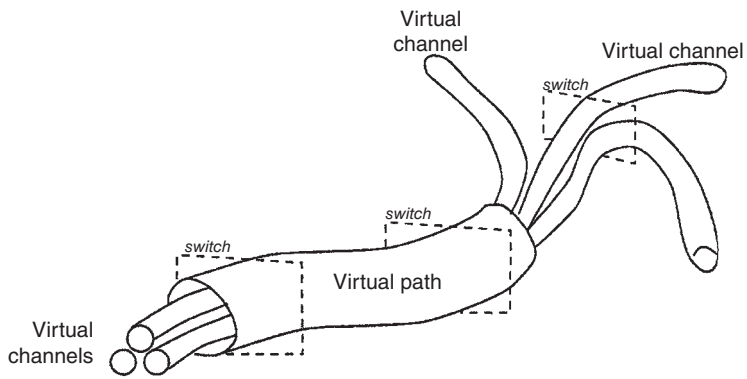


Figure 19.6 Virtual channels and virtual paths.

stored as values in the switch look-up tables, forming a linked list of table entries along the path (Fig. 19.12). A connection-establishment procedure is responsible for setting up the proper look-up table values whenever a new channel or path is initiated.

Depending on the kind of switch, either just the VPI information is used in routing or both VPI and VCI fields are used. Because logically VCs are viewed as being contained inside VPs, VP-only switches are not allowed to substitute VCI fields of outgoing cells; VP-only routing is considered a lower sublayer than VC routing [3].

Note also that cells entering a particular switch destined for different output ports must have distinct VPI values. Within a given VP, different VCs are distinguished by differing VCI values. However, different channels may share the same VCI value if their VPI fields differ (they belong to a different virtual path).

19.4. CLASSICAL IP OVER ATM

Classical IP over ATM (RFC 1577) [13, 14] was proposed by the Internet Engineering Task Force (IETF) as a way of connecting IP-based workstations on ATM. RFC 1577 basically emulates the IP layer (network layer) over ATM to provide end-to-end connectivity to the higher layers. In this approach, the IP end stations connected to the ATM cloud are divided into logical IP subnets (LISs). The subnets are administered in the same manner as the conventional subnets. Hosts connected to the same subnetwork can communicate directly. However, communication between two hosts on different LISs is only possible through an IP router, regardless of whether direct ATM connectivity is possible between the two hosts.

Implementation of classical IP over ATM requires a mapping between IP and ATM addresses. IP addresses are resolved to ATM addresses using the ATM address resolution protocol (ATMARP) and vice versa using the inverse ATMARP (InATMARP) within a subnet. ATMARP is used for finding the ATM address of a device given the IP address. It is analogous to the IP-ARP associated with IP protocol. Just like conventional ARP, it has a quintuple associated with it: source IP address, source ATM address, destination IP address, and destination ATM address. On the other hand, InATMARP is used to find the TP address of a station given the ATM address [almost equivalent to conventional reverse address resolution protocol (RARP)]. Typical use of InATMARP is by the ATMARP server to find out the IP address of the station connected to the other end of an open virtual circuit. This information is used to update database entries and to refresh the entry on time-outs.

Every end station is configured statically with the address of the ATMARP server. On initialization, the end station opens a virtual channel connection (VCC) to the ATMARP server. The ATMARP server, on detecting a new VCC, performs

an InATMARP on it to find the IP address of the end station connected at the other end of the VCC. This information is stored in the tables of the ATMARP server for further use. Each end station maintains a local ARP cache that acts as the primary cache, and the APR server acts as a secondary cache. If the ATM end station wants to contact another station, it will query its local cache first for the ATM address for a given IP address. If that fails, it queries the ARP server for the ATM address. Once it has the ATM address of the destination, the end station proceeds to open a direct VCC to the destination.

The basic drawback to this approach is that it works only for IP-type traffic; it does not support multicast or broadcast. It requires static assignment of the ARP server address, and the ARP server becomes the single point of failure. The IETF has recently removed some of these drawbacks by introducing a new concept to enhance RFC 1577, that is, multicast address resolution server. Work on it has been ongoing since late 1994.

19.5. ATM LAN EMULATION

LAN emulation (LANE) [15–17] has been proposed by ATM Forum and has been widely accepted by the ATM industry as a way to emulate conventional LANs. The necessity of defining LANE arose because most of the existing customer premises networks use LANs such as IEEE 802.3/802.5 (Ethernet and Token Ring) and customers expect to keep using existing LAN applications as they migrate toward ATM. To use the vast repertoire of LAN application software, it became necessary to define a service on ATM that could emulate LANs. The idea is that the traditional end-system applications should interact as if they are connected to traditional LANs. This service should also allow the traditional (legacy) LANs to interconnect to ATM networks using today's bridging methods.

LANE has been defined as a MAC service emulation, including encapsulation of MAC frames (user data frames). This approach, as per ATM Forum, provides support for maximum number of existing applications. This is not easy because there are some key differences between legacy LANs and ATM networks.

The main objective of LAN emulation service is to enable existing applications to access an ATM network via protocol stacks, such as NetBIOS, IPX, IP, and AppleTalk, as if they were running over traditional LANs. In many cases, there is a need to configure multiple separate domains within a single network. This objective is fulfilled by defining an emulated LAN (ELAN) that comprises a group of ATM-attached devices. It appears as a group of stations connected to a IEEE 802.3 or 802.5 LAN segment. Several ELANs could be configured, and membership in an ELAN is independent of the physical location of the end station. An end station could belong to multiple ELANs.

19.5.1. Components

LANE has four basic components: LAN emulation client (LEC), LAN emulation configuration server (LECS), LAN emulation server (LES), and broadcast and unknown server (BUS).

LEC is an entity in the ATM workstation or ATM bridges that performs data forwarding, address resolution, and other control functions. This provides a MAC-level emulated Ethernet/IEEE 802.3 or BEE 802.5 service interface to applications running on top. It implements the LANE usernetwork interface (LUNI) when communicating with other entities within the emulated LAN.

The LES implements the control coordinating function for the ELAN. It provides a facility for registering and resolving MAC addresses or route descriptors to ATM addresses. LECs register the LAN destinations they represent with the LES. A client can also query the LES when the client wishes to resolve a MAC address to an ATM address. A LES will either respond directly or forward the query to other clients so they may respond.

BUS handles data sent by an LEC to the broadcast MAC address ("F-" hex), all multicast traffic, and, as an option, some initial unicast frames sent before the target ATM address is resolved.

19.5.2. LEC Initialization Phases

The basic states that a LEC goes through before it is operational are shown in Fig. 19.5 and described as follows:

Initial state: In this state LES and LEC know certain parameters (such as address, ELAN name, maximum frame size) about themselves.

LECS connect phase: LEC sets up a call to LECS. The VCC that is opened is referred to as configuration-direct VCC. At the end of configuration, this VCC may be closed by the LEC.

Configuration phase: LEC discovers LES in preparation for join phase.

Join phase: During the join phase, LEC establishes its control connections to the LES. Once this phase is complete, the LEC has been assigned a unique LEC identifier (LECID), knows the emulated LAN's maximum frame size and its LAN type, and has established the control VCC with the LES.

Initial registration: After joining, an LEC may register any number of MAC addresses in addition to the one registered during the join phase.

BUS connect: In this phase a connection is set up to the BUS. The address of the BUS is found by issuing an *LE-ARP* for an ATM address with all 1s. The BUS then establishes a multicast-forward VCC to the LEC.

19.5.3. Connections

An LEC has separate VCCs for control traffic and for data traffic. Each VCC carries traffic for only one ELAN. The VCCs form a mesh of connections between the LECs and other LANE components such as LECS, LES, and BUS.

19.5.4. Control Connections

A control VCC links the LEC to the LECS and LEC to the LES. The control VCCs never carry data frames and are set up as a part of the LEC initialization phase. The control connection terminology is as follows:

Configuration-direct VCC is a bidirectional point-to-point VCC set up by a LEC as part of the LECS connect phase and is used to obtain configuration information, including the address of LES. This connection may be closed after this phase is over.

Control-direct VCC is a bidirectional point-to-point VCC *to* the LES set up by a LEC for sending control traffic. This is set up during the initialization process. Because LES has the option of using the return path to send control data to the LEC, this requires the LEC to accept control traffic on this VCC. This VCC must be maintained open by both LES and LEC while participating in the ELAN.

Control-distribute VCC is a unidirectional point-to-multipoint or point-to-point VCC from LES to the LEC to distribute control traffic. This is optional, and LES, at its discretion, may or may not set this up. This VCC is also set up during the initialization phase. This VCC, if set up, must be maintained while participating in the ELAN.

19.5.5. Data Connections

Data VCCs connect the LECs to each other and to the BUS. These carry either Ethernet or Token Ring data frames and under special conditions a flush message (optional). Apart from flush messages, data VCCs never carry control traffic:

Data-direct VCC is a bidirectional point-to-point VCC established between LECs that want to exchange unicast data traffic.

Multicast-send VCC is a bidirectional point-to-point VCC from LEC to BUS. It is used for sending multicast data to the BUS and for sending initial unicast data. The BUS may use the return path on this VCC to send data to the LEC, so this requires the LEC to accept traffic from this VCC. The LEC must maintain this VCC while participating in the ELAN.

Multicast-forward VCC is either a point-to-multipoint VCC or a unidirectional point-to-point VCC from the BUS to the LEC after the LEC sets up a

multicast-send VCC. It is used for distributing data from the bus. The LEC must attempt to maintain this VCC while participating in the ELAN.

19.5.6. Operation

To get to the operational state, that is, the state at which the LEC can start exchanging information with other LECs, it has to go through an initialization process that consists of several phases. First, if required, it must contact the LECS. This phase is optional and may not exist if a preconfigured switched virtual circuit or permanent virtual circuit (PVC) to LES is used. The LEC will locate the LECS by using the following mechanisms to be tried in the following order: (i) Get the LECS address via interim local management interface (ILMI) using the ILMI Get or ILMI Get Next to obtain the ATM address of the LECS for the UNI; (ii) using the well-known LECS address: If LECS address cannot be obtained via ILMI or if LEC is unable to establish a configuration direct VCC to that address, then an ATM Forum specified well-known address "47.00.79.00.00.00.00.00.00.00.00.00-00.A0.3E. 00.00.01-00" hex must be used to open a configuration direct VCC; (iii) using a well-known PVC: If VCC could not be established to the well-known address in the previous step then the well-known PVC of virtual path identifier = 0 and virtual channel identifier = 17 (decimal) must be used.

The configuration phase prepares the LEC for the join phase by providing the necessary operating parameters for the emulated LAN that the LEC will join. Once the LECS is found, then LEC sends a *LE_Configure_Request* and waits for a *LE_Configure_Response,* which is a part of the LE configuration protocol. All control frames have the structure shown in Fig. 19.7. Marker is always a fixed 2-byte value "FFOO" hex. The op-code determines the type of control frame, for example, "0001" for *LE_Configure_Request* "0101" for *LE_Configure_Response*, etc. Status is used in the responses to inform about reasons of denial for the requests or to indicate success. Type-length values are used to exchange specific information in the control frames such as timer values and retry counts.

During the configuration, the LECS provides the LEC with the ATM address of LES and also provides all kinds of timers values, time-out periods, and retry counts.

Armed with this information, LEC enters the join phase. Here, the LEC establishes its connection with the LES and determines the operating parameters of the emulated LAN. The LEC implicitly registers one MAC address with the LES as a part of the joining process. LEC must initiate the UNI signaling to establish a control-direct VCC (or use a control-direct PVC) and then send a *LE_JOIN_Request* over this VCC to the LES. The LES may optionally establish a control-distribute VCC back to the LEC.

After that the LES will send back a *LE_JOIN_Response* that may be sent on either control direct or control distribute (if created). To each LEC that joins, the LES assigns a unique LECID.

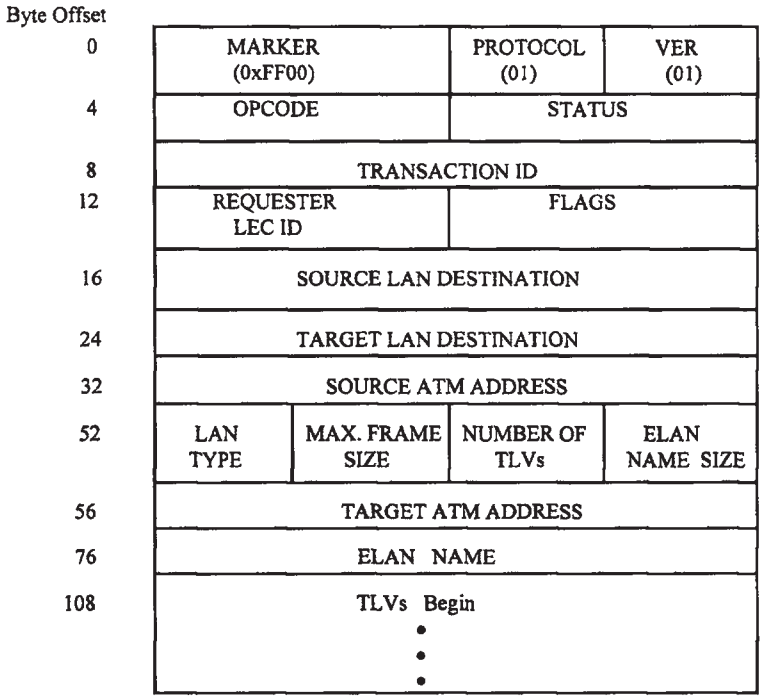


Figure 19.7 LANE frame format [15]. Copyright 1995 The ATM Forum.

If the join phase is successful, then the LEC is allowed to register additional MAC addresses, which it represents with the LES. This is called the registration phase. However, this can happen any time and is not restricted to this phase. However, additional registrations cannot be done before joining the ELAN.

This is followed by the BUS connect phase in which LEC has to establish connection to the BUS. For this purpose, the LEC needs to find out the address of the BUS. This is accomplished by the ARP. In this procedure whenever a LEC is presented with a frame for transmission whose LAN destination is unknown to the client, it must issue LANE ARP (*LE_ARP*) request frames to the LES over its control-direct VCC. The LES may issue an LE-ARP reply on behalf of a LEC that had registered the LAN destination earlier with the LES or alternatively can forward the request to the appropriate client(s) using the control-distribute VCC or one or more control-direct VCCs, and then the *LE_ARP_Reply* from the appropriate LEC will be relayed back over the control VCCs to the original requester. Each LEC also maintains a local cache of addresses.

For connecting to the BUS, LEC first issues an *LE_ARP_Request* to the LES for the broadcast MAC address, that is, all 1s-("FFFFFFFFFFFF" hex). The

LE_ARP_Response gives the ATM address of the BUS. The LEC then proceeds to set up a multicast-send VCC to the BUS, which then immediately opens a multicast-forward VCC back to the LEC. At this point the LEC is considered operational.

Now, if the LEC wants to exchange information with another LEC, it can use the address resolution procedure to get its address and then set up a data-direct VCC to the other LEC and transfer information. However, to save time, if the target LEC's address is not known, then the originating LEC issues a *LE_ARP_Request* and starts sending frames through the BUS. Once the *LE_ARP_Reply* is received, the LEC is required to stop using the BUS and open a data-direct VCC. Despite all this, ATM guarantees in-order delivery and therefore a flush message is sent to BUS that ensures that no frames are transmitted on the data-direct VCC until all the previous ones routed through the BUS are delivered. Flush request message is a way to inform the other side that following that request, data will be transmitted on a different channel, for example, switching from multicastsend to data-direct VCC. The flush request needs to be responded by flush response so that the side issuing the flush request understands that all the previously sent messages have been delivered on the old channel and it is safe to switch channels and still maintain in-order delivery of messages.

19.6. GFP AND LCAS

While long-haul communications networks are currently dominated by SONET/SDH, a wide range of data center protocols may also have applications for long-distance transport, including ESCON, FICON, Fibre Channel, Ethernet, and some nonstandard protocols such as those used in a Parallel Sysplex (see Chapters 17, 20, 21, and 22). Until recently, SONET/SDH networks were optimized for time-division multiplexed traffic that could be classified into predictable, well-defined incremental bit rates (characteristic of voice traffic). With the recent growth in data center traffic, these networks now face the challenge of handling less predictable, bursty traffic with variable bandwidth utilization. This has led to the development of new standards intended to extend the useful lifetime of SONET/SDH networks and leverage the low cost and established management, installation, and service expertise surrounding these networks.

The International Telecommunications Union (ITU) has recently proposed a new industry standard G.7041 called Generic Frame Procedure (GFP) (additional information on this approach is provided in Chapter 15). This is intended to allow the mapping of higher layer client signals in a variety of different protocols into a frame structure compliant with SONET/SDH so that this traffic can be carried over a common transport network. The client signals include standard datacom protocols with 8B/10B data encoding, such as Fibre Channel, FICON, and ESCON, or protocol data units (PDUs) such as IP or Ethernet traffic. Since there

is a large amount of SONET infrastructure in use by telecom carriers and other service providers, GFP is seen as the means to allow enterprise systems to carry data traffic over existing SONET networks at very low incremental cost. In turn, this enables channel extensions over hundreds or thousands of km for applications such as disaster recovery. In this regard, GFP has also been implemented as part of many WDM platforms for dark fiber applications (see Chapter 15).

There are two modes of operation for these systems. GFP-Framed (GFP-F) maps each client frame into a single GFP frame and should be used when the client signal is framed by the client protocol. For example, GFP-F can encapsulate complete Ethernet frames with a GFP header. This packet-oriented approach is generally optimized for bandwidth efficiency, at the expense of latency. By contrast, GFP-transparent (GFP-T) allows more efficient transport of low-latency protocols by the mapping of multiple 8B/10B encoded client data streams into a common block of 64B/66B encoded data for transport within a GFP frame. In this character-oriented mode, instead of buffering an entire client frame and then encapsulating it into a GFP frame, the individual characters of the client data stream are extracted, and a fixed number of them are mapped into periodic fixed-length GFP frames. This mapping occurs regardless of whether the client character is a data or control character, thus preserving the client 8B/10B control codes. It is still possible to perform frame multiplexing with GFP-T, if desired. Both approaches include basic functions such as frame delineation, client multiplexing, and encapsulation compliant with network switching and routing functions.

As shown in Fig. 19.8, a GFP frame consists of a core header, a payload header, an optional extension header, the GFP payload, and an optional frame check sequence (FCS). The core header is 4 bytes long and consists of two fields: a 2-byte payload length indicator (PLI), which indicates the size of the core header in bytes, and a 2-byte core header error correction field (cHEC), which is a cyclic redundancy check (CRC) on the core header. The payload field, of course, contains the client data mapped as either GFP-F or GFP-T, and the FCS ensures the integrity of the frame. Both the core header and payload are scrambled to ensure an adequate number of transitions between 1 and 0 bits to enable adequate clock recovery (this is the only way that the receiver can remain synchronized with the transmitter). The variable-length payload header consists of a payload type field and a type Header Error Correction (tHEC) field (optionally, the payload header may include an extension header, which we will not describe in detail). The payload type field consists of several subfields:

- The Payload Type Identifier (PTI) subfield identifies the type of frame. Two values are currently defined: user data frames and client management frames.
- The Payload FCS Indicator (PFI) subfield indicates the presence or absence of the payload FCS field.

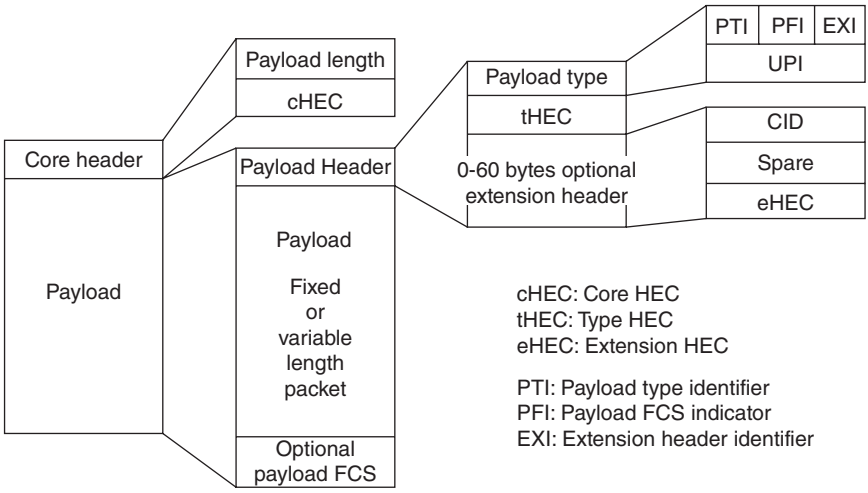


Figure 19.8 GFP frame structure.

- The Extension Header Identifier (EXI) subfield identifies the type of extension header in the GFP frame. Extension headers facilitate the adoption of GFP for different client-specific protocols and networks. Three kinds of extension headers are currently defined: a null extension header, a linear extension header for point-to-point networks, and a ring extension header for ring networks.
- The User Payload Identifier (UPI) subfield identifies the type of payload in the GFP frame. The UPI is set according to the transported client signal type. Currently defined UPI values include Ethernet, point-to-point protocols including IP and MPLS, Fiber Channel, FICON, ESCON, and Gigabit Ethernet.

There are two basic types of GFP frames: client and control frames. Control frames (also known as idle frames) consist of a core header field only with no payload data; they are used to compensate for gaps between lower speed client signals being mapped onto a higher speed transport link. Client frames can be further classified as either client data frames (used to transport client data) or client management frames (used to transport management information such as loss of signal). The two types of client frames can be distinguished based on their payload type indicators. Client frames are given priority over management frames when multiplexing data.

The basic GFP-T procedure for mapping protocols such as ESCON or FICON involves decoding each 10-bit character of an 8B/10B data sequence, and mapping the result into either an 8-bit data character or a recognized control character.

This data is then re-encoded as a 64B/65B data sequence, with control characters mapped into a predetermined set of 64/65B control characters. In GFP terminology, the resulting data sequences or control characters are known as words (this differs from the server definition of a word, which is usually taken as either a 4-byte quantity or a 40-bit string of four 8B/10B characters. We will use the GFP terminology for consistency throughout the remainder of this discussion). A group of 8 such words is assembled into an octet, which is provided with additional control and error flags (note that this differs from the server definition of an octet, which is usually taken as an 8-bit byte). A group of 8 octets is then assembled into a "superblock," scrambled, and a CRC error check field is added. The resulting frames are compliant with routing through a SONET/SDH network flow control, including quality of service and related features. By reversing this process, the original 8/10 encoded data is reassembled at the other end of the network.

The ITU standard G.707/Y.1332 defines virtual concatenation (VC), a technique that allows SONET/SDH circuits to be grouped into arbitrarily sized bandwidth increments for more efficient transport of client protocols. The channel bandwidth is divided into smaller individual containers, which are grouped together and logically represented by a virtual concatenation group (VCG). The members of a VCG can be routed independently over an existing SONET/SDH network (by simply upgrading the network end points). The containers can take different paths through the network and incur different propagation delays; the destination receiver stores containers as they arrive and reassembles the desired data stream.

A related ITU standard and further enhancement of VC, G.7042, defines a method for dynamically increasing or decreasing the bandwidth capacity of virtual channels such as TDM containers over a SONET/SDH network. This is known as the Link Capacity Adjustment Scheme (LCAS). The intent is to provide flexible bandwidth-on-demand allocations for data center protocols when operating over SONET/SDH networks, as opposed to the conventional telecom provisioning schemes, which require some *a priori* nominal definition of channel bandwidth capacity. Since data traffic may come in bursts and generally is less predictable than voice traffic, it can be inefficient and expensive to provision network bandwidth based on estimated or peak usage. LCAS is intended to help address this problem by allocating bandwidth in a more flexible fashion, responding to network traffic loads in near real time. LCAS is also useful for load balancing across different network paths and managing quality of service; it enables carriers to oversubscribe the network and still remain profitable through tiered service-level agreements (SLAs) for data services. LCAS is also intended to improve fault recovery by providing so-called hitless upgrades, meaning that data traffic continues to flow uninterrupted while the network equipment is changing the bandwidth capacity of the transport media. In addition, failed members

in a virtual concatenation group can be removed by LCAS in a hitless fashion; the network bandwidth decreases automatically when a member fails and is automatically restored when the member is repaired. When combined with diverse path routing, this function is intended to increase the survivability of network traffic without requiring the additional expense of allocating network bandwidth just for protection purposes. The combination of GFP and LCAS is intended to extend the usable lifetime of the installed SONET/SDH network infrastructure and to accommodate the growing amount of data traffic using these networks.

REFERENCES

1. Hac, A., and H. B. Mutlu. 1989, November. Synchronous optical network and broadband ISDN protocols. *Computer* 22(11):26–34.
2. Stallings, W. 1992. *ISDN and broadband ISDN*, 2nd ed. New York: Macmillan.
3. Minoli, D. 1993. *Enterprise networking: Fractional TI to SONET frame relay to BISDN*. Boston: Artech House.
4. DeCusatis, C. 1995. Data processing systems for optoelectronics. In *Optoelectronics for data communication*, eds. R. Lasky, U. Osterberg, and D. Stigliani, Chapter 6. New York: Academic Press.
5. Miller, A. 1994, June. From here to ATM. *IEEE Spectrum* 31(6):2&24.
6. Jungkok Bae, J., and T. Suda. 1991, February. Survey of traffic control schemes and protocols in ATM networks. *Proc. IEEE* 79(2):170–189.
7. Roohalamini, R., V. Cherkassky, and M. Garver. 1994, April. Finding the right ATM switch for the market. *Computer* 27(4):16–28.
8. ATM Forum Inc. 1995. *ATM user network interface (UNI) specification Version 3.1*, *l/e*. New York: Prentice Hall.
9. The ATM Forum Technical Committee. 1994, September. DS 1 physical layer specification. Technical Report AF-PHY-0016.000, The ATM Forum.
10. The ATM Forum Technical Committee. 1995, November 7. Physical interface specification for 25.6 Mb/s over twisted pair cable. Technical Report AF-PHY-0040.000, The ATM Forum.
11. The ATM Forum Technical Committee. 1996, January. 622.08 Mb/s physical layer specification. Technical Report AF-PHY-0046.000, The ATM Forum.
12. Klessig, B. 1995, July. Status of ATM specifications. http://www.3corn.com/ Ofiles/mktglpubs/ 3tech/795atmst.html.
13. 13.RFC 1577. Classical IP over ATM, request for comments. Internet Engineering Task Force.
14. Comer, D. E. 1995. *Internetworking with TCP/IP—Volume 1*, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.
15. ATM Forum. 1995. *LAN emulation over ATM Version 1.0*, af-lane-0021.000. ATM Forum, 303 Vintage Park Drive, Foster City, Calif.
16. Siu, K., and R. Jain. 1995. A brief overview of ATM: Protocol layers, LAN emulation and traffic management. ACM SIGCOMM *Comput. Commun. Rev.* 25(2):6–20.
17. Finn, N., and T. Mason. 1996. ATM LAN emulation. *IEEE Commun. Mag.* 34(6):96–100.

Case Study Facilities-Based Carrier Network Convergence and Bandwidth on Demand

Courtesy of Cisco Systems

Application: A global network service provider redesigns its core network to enable the convergence of voice and data traffic, while providing scalable bandwidth on demand services.

Description: The rapidly growing customer base for telecom and datacom services is also demanding higher reliability and larger bandwidths for enterprise customers. Many large enterprises are deploying fault-tolerant access networks with 20–30 Gbit/s aggregate bandwidth, for applications including 10 Gbit/s Ethernet LANs, 8–10 Gbit/s Fibre Channel and FICON connections, managed SONET/SDH, and digital video services that cannot be easily delivered over SONET. Furthermore, these customers demand dynamic bandwidth reprovisioning, including the ability to change services (OC-n to Gigabit Ethernet, for example) within 4 hours of a request with no discernible down time. The carrier providing such services addressed this challenge by deploying two separate point-of-presence (POP) locations with redundant hardware, then provisioning a virtual carrier network service for each customer. Protected OC-12, -48, or -192 rings were installed between customer sites and the POPs, so that an assortment of Ethernet, storage, and video connections could be deployed as required. The infrastructure was based on an optical WDM platform (the Cisco ONS 15454 MSPP/MSTP), which offered the ability to upgrade enterprise customers to a Geographically Dispersed Parallel Sysplex (GDPS) in the future. The network includes 64 wavelengths of protected traffic and reconfigurable optical add/drop multiplexing (ROADM) capability to meet reconfiguration needs. The WDM platform enabled end-to-end wavelength path provisioning (similar to SONET service provisioning) from a central location, including automatic power control

for the optical amplifiers (note that amplifiers were not required for distances up to 80 miles, which allowed the extension of the network over previously unused dark fiber in some areas). For maximum flexibility and minimum sparing cost, the design includes wavelength tunable lasers adjustable over 4 channels of the ITU grid C-band. These features allow new wavelengths to be provisioned for service within a few hours instead of days. The same network accommodates SONET services with network equipment blades supporting up to 12 SFP transceivers in any combination of OC-3, OC-12, and OC-48 line rates on a port-by-port basis. The SFPs are not installed until needed, which supports a pay-as-you-grow business model and makes it possible to reprovision optical ports for optimal efficiency. The convergence of SONET and storage traffic, combined with an agile network that can be quickly reconfigured in response to changing traffic conditions, provides one example of how emerging optical networking technologies can be combined to provide business value.

20

Fibre Channel—The Storage Interconnect

Scott Kipp
Brocade Corporation

Alan Benner
IBM Corporation

This chapter discusses the evolution of storage area networks (SANs) and their most common underlying protocol—Fibre Channel. While laptops and desktops connect to local area networks (LANs) via the Ethernet protocol and physical layer, servers and mainframes connect to SANs via the Fibre Channel protocol and physical layer. Fibre Channel creates a fabric operating at multiple gigabit-per-second (Gbps) and interconnects servers to a variety of storage devices. The SAN requires links with more bandwidth over shorter distances than most LAN connections and was the first widely deployed application for Gbps multimode fiber. The English spelling "Fibre" was used to convey that this standard supports both optical fiber and copper links. This chapter concentrates on the application of fiber optics to the physical layer of the SAN.

While telecom networks ran at Gbps speeds over single-mode fiber, Fibre Channel was among the first to use vertical cavity surface-emitting lasers (VCSELs) over multimode fiber. To keep Fibre Channel from becoming an expensive niche technology, the Technical Committee T11 of the American National Standards Institute (ANSI) defined Gbps Fibre Channel links affordably to increase adoption. The low cost and high speed of the Fibre Channel physical layer is one of the main reasons for its continued success over competing technologies. From 850-nanometer (nm) VCSELs over multimode fiber to 1550-nm distributed feedback (DFB) lasers over single-mode fiber, Fibre Channel has kept pace with advances in storage capacity that continually grows at annual rates even faster than Moore's Law.[1] The SAN will continue to be the most data intensive aspect of enterprise networks, and Fibre Channel is its foundation.

[1]http://www.wired.com/wired/archive/14.10/cloudware_pr.html, The Information Factories, George Gilder.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

20.1. INTRODUCTION TO FIBRE CHANNEL

In the early 1990s, the enterprise storage industry began seeing the limitations of the parallel bus technology used to connect disk drives to servers and mainframes. Every time the speed of the bus doubled, the supported distance of the bus was typically cut in half. Since data centers were continually growing in speed and scale, the storage community needed a solution that increased speed and distance. The mainframe computing industry initially developed the serial Enterprise System Connection (ESCON) to overcome the limitations of the shared parallel bus (see Chapter 21). A serializer/deserializer (SERDES) is the essential component of converting the parallel electrical signals into a high-speed serial signal as shown in Fig. 20.1. The serializer creates a high-speed serial signal and drives a transceiver that performs the electrical to optical (E/O) conversion (see Chapter 5). The high-speed optical signal travels much longer distances because of the high-bandwidth-distance product of the fiber. The transceiver performs optical to electrical (O/E) conversion and feeds the deserializer that creates a slower parallel electrical signal.

The open systems community that used the parallel small computer systems interface (SCSI) bus wanted the same benefits of longer distances and higher speeds. The combined interests of the storage community found a home in the T11 Technical Committee that continues to define Fibre Channel interfaces. The committee made a crucial decision to standardize 850-nm VCSELs using multimode fiber to achieve links up to 500 meters.

The use of low-cost VCSELs was a big improvement in cost over previous single-mode solutions. Most fiber-optic solutions before Fibre Channel used

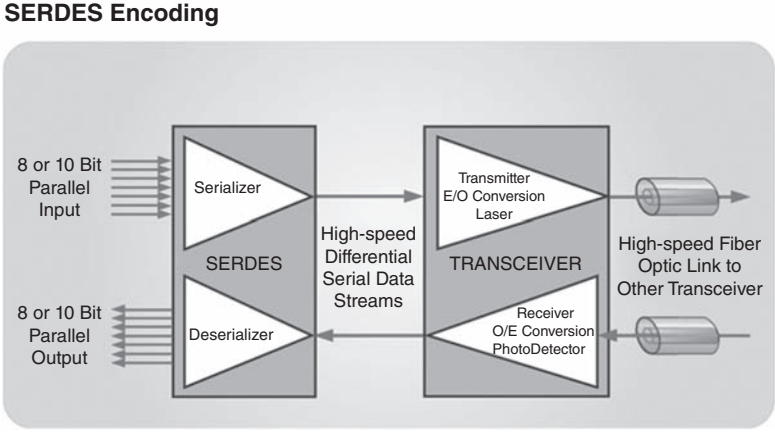


Figure 20.1 Low-speed parallel electrical signals from a printed circuit board are fed into the SERDES that encodes the bits into a high-speed serial stream.

either light-emitting diodes (LEDs), Fabry-Perot lasers, or DFB lasers. LEDs could only be modulated to a few hundred megabits-per-second and had high failure rates. Fabry-Perot and DFB lasers were expensive and did not lend themselves to high-volume manufacturing. VCSELs were a better technology choice since they could be tested in wafer form and easily packaged. Fibre Channel was the first application with wide adoption of 850-nm VCSEL technology.

The first standardized, pluggable transceivers to meet the needs of the SAN community were known as gigabit interface converters or GBICs. The term *GBICs* is still commonly used for transceivers, but with the creation of the small form factor pluggable (SFP) transceiver in 2000, the industry quickly converted to SFPs. For applications with fixed optics that were soldered to the board, 1X9 pin transceivers were used first and were comparable to GBICs. These were replaced in a similar manner to the GBICs by the small form factor (SFF) transceiver. These transceivers are based on standard electrical and housing interfaces and support either optical or copper solutions as seen in Fig. 20.2; for more details on transceiver packages, see Chapter 11.

Fibre Channel links are defined for a variety of optical fibers and copper cables. Short-distance implementations could use Category 5 ot 6 (CAT-5, CAT-6) cables or twin-axial cables. The vast majority of initial deployments of Fibre Channel links used OM2 fiber (see Chapter 2), while more recent deployments use OM3 fiber. Details about supported distances will be provided later in this chapter; distances for 1 gigabit links are summarized in Fig. 20.3. The specification of 1 Gigabit Fibre Channel (1GFC) links was later used by the Institute of Electrical and Electronic Engineers (IEEE) as the basis for some of the optical links defined by Gigabit Ethernet.

Logically, Fibre Channel is structured as a set of hierarchical functions as shown in Fig. 20.4. The lowest level or FC-0 describes the physical interface, including transmission media and transceivers that can operate at various data rates. The FC-1 layer describes the 8B/10B transmission code used to provide DC balance for the transmitted bit stream, to separate control bytes from data bytes, to simplify bit, byte, and word alignment, and to detect some types of transmission and reception errors. The FC-2 layer (FC-FS-2)[2] is the signaling protocol, perhaps the most complex layer, which specifies the rules needed to transfer blocks of data, classes of service, packetization, sequencing, error detection, segmentation, reassembly, and other services. The FC-3 layer provides services that are common across multiple ports of a network node. The FC-4 layer maps preexisting upper level protocols (ULPs) such as Internet Protocol (IP),

[2]http://www.t11.org/t11/docreg.nsf/ufile/06-085v3, Fibre Channel—Framing and Signaling-2 (FC-FS-2), Robert Nixon.

Fibre Channel Transceivers

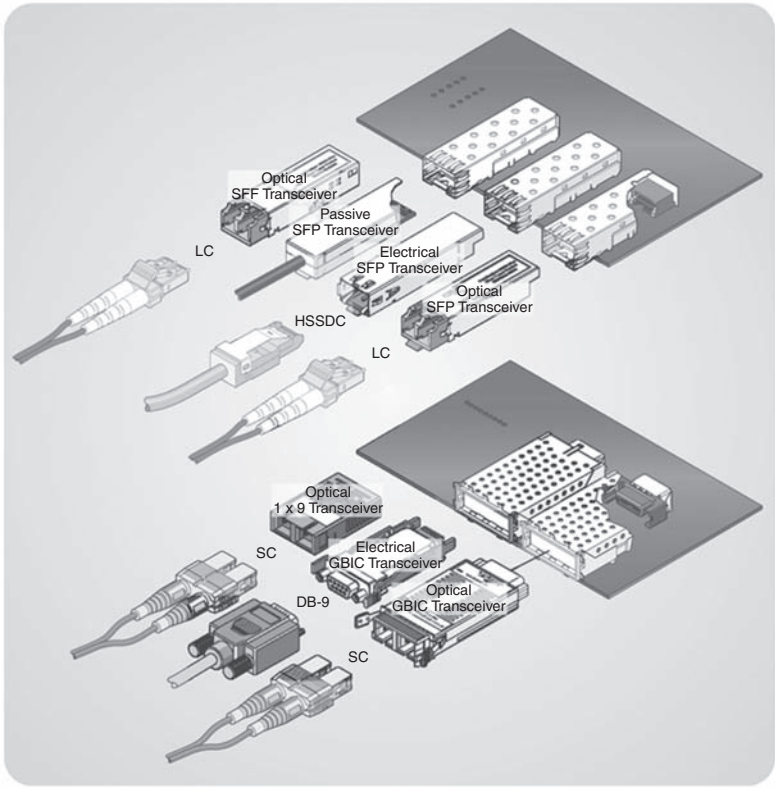


Figure 20.2 The first generation of transceivers used in Fibre Channel is shown on the bottom of this figure, while their replacements are shown in the upper half.

FICON's Single Byte Command Code Set (SBCCS), or the Small Computer Systems Interface (SCSI) to the Fibre Channel layers.

A Fibre Channel network is made up of one or more bidirectional point-to-point serial data channels. Physically, this network can be set up in several different topologies: (1) a single point-to-point link between two ports called N_Ports, (2) a network of multiple N_ports, each linked into a switching fabric through an F_port, or (3) a ring topology called an Arbitrated Loop, which allows multiple N_port connections without switch elements. Fibre Channel–Arbitrated Loop (FC-AL) is a sharing topology—a single port, called an NL_port, arbitrates for access to the entire loop and prevents access by any other NL_ports while it is communicating. Each N_port resides on a computer, disk drive, or other piece of hardware called a node. A single Fibre Channel node implementing one or

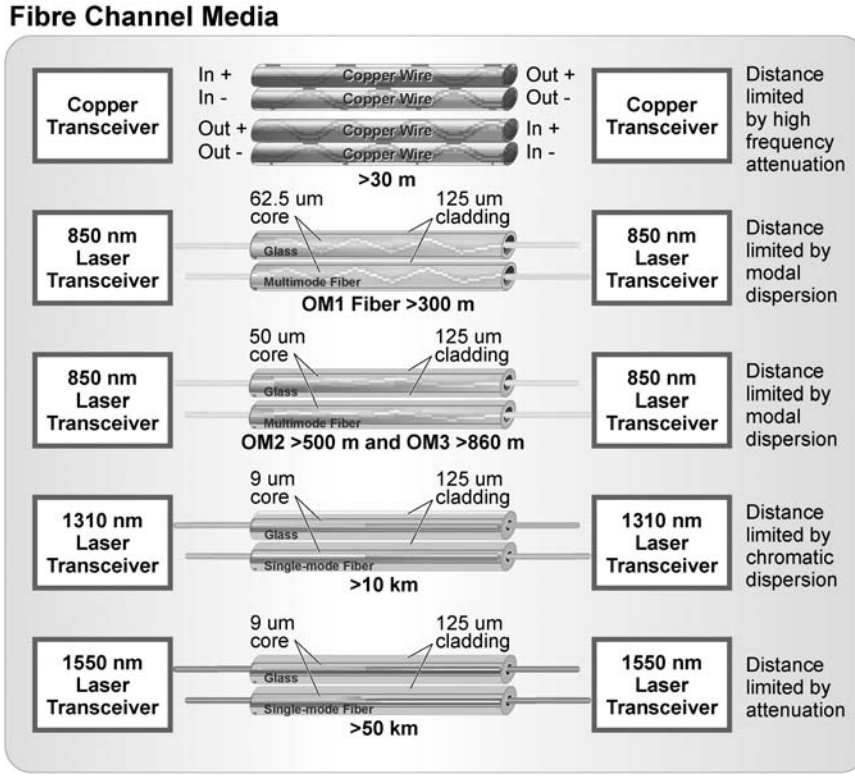


Figure 20.3 Unrepeated distances for Fibre Channel defined links over multiple media types at 1-Gigabit Fibre Channel (1GFC).

more N_ports provides a bidirectional link with FC-0 through FC-2 or FC-4 layer services through each N_port. Fibre Channel also defines several other types of ports.

An example of another Port type is the Expansion Port (E_Port) that interconnects multiple switches through interswitch links (ISLs). When an E_Ports is attached to another switch's E_Port, the switches form a fabric that behaves like one large switch. The protocols to form a fabric are defined in Fibre Channel–Switch Fabric (FC-SW-4). Many vendor switches incorporate additional, proprietary features such as data compression or link aggregation. Proprietary features make the selection of a particular brand of switch very important. Some switches may be configured to reduce latency by not storing frames before forwarding them (so called cut-through switches). Fibre Channel fabrics create intelligent networks that control the access to storage devices.

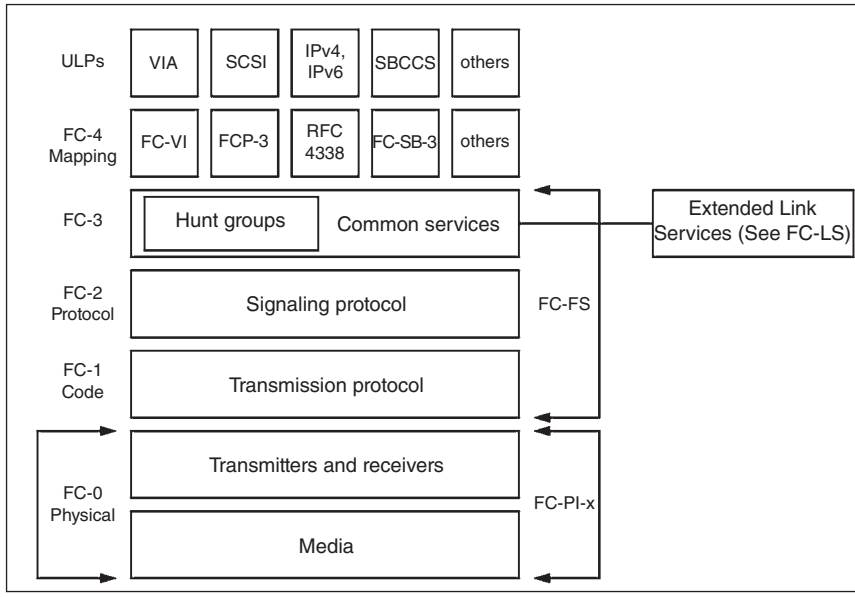


Figure 20.4 The Fibre Channel layered architecture.

The fabric may be a mix of switched links and arbitrated loop technologies; a fabric port capable of operating on a loop is called an FL_port. The standard also defines a G_port, which may function as either an E_port or an F_port depending on how it is connected, and a GL_port, which can operate as either an F_port, an E_port, or an FL_port. Fibre Channel functions are topology independent and rely on a series of "login" procedures to determine the topology of the network to which it is connected.

20.1.1. Fibre Channel Data Rates

The maximum data transfer bandwidth over a link depends both on physical parameters, such as clock rate and maximum baud rate, and on protocol parameters, such as signaling overhead and control overhead. The data transfer bandwidth can also depend on the communication model, which describes the amount of data being sent in each direction at any particular time.

The primary factor affecting communications bandwidth is the clock rate of data transfer. The base clock rate for data transfer under 1GFC is 1.0625 GHz, with 1 bit transmitted every clock cycle. Higher rate links are also defined, including double-(2GFC), quadruple-(4GFC), and 8GFC speed links. Higher data rates are designed to autonegotiate to the lowest supported link rate to facilitate backward compatibility, with the exception of 10 Gbit/s (10GFC) links which have

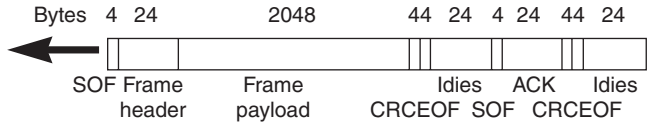


Figure 20.5 Typical Fibre Channel data frame.

only been used as ISLs. At this time, 4 Gbit/s links are commonly in use, and early adoption of 8 Gbit/s links is expected to occur in 2008.

Figure 20.5 shows a sample communication model for calculating the achievable data transfer bandwidth over a link. The figure shows a single Fibre Channel Frame, with a payload size of 2048 bytes. To transfer this payload and an acknowledgment, the following overhead elements are required:

SOF: Start of Frame delimiter, for marking the beginning of the Frame (4 bytes)

Frame Header: Indicating source, destination, sequence number, and other Frame information (24 bytes)

CRC: Cyclic Redundancy Code word, for detecting transmission errors (4 bytes)

EOF: End of Frame delimiter, for marking the end of the Frame (4 bytes)

Idles: Inter-Frame space for error detection, synchronization, and insertion of low-level acknowledgments (24 bytes)

ACK: Acknowledgment for a Frame from the opposite Port, needed for bidirectional transmission (36 bytes)

Idles: Inter-Frame space between the ACK and the following Frame (24 bytes)

The sum of overhead bytes in this bidirectional transmission case is 120 bytes, yielding an effective data transfer rate of 100.369 MBps:

$$1.0625[\text{Gbps}] \times \frac{2048[\text{payload}]}{2168[\text{payload} + \text{overhead}]} \times \frac{1[\text{byte}]}{10[\text{codebits}]} = 100.369$$

Thus, the full-speed link provides better than 100 MBps data transport bandwidth, even with signaling overhead and acknowledgments. The achieved bandwidth during unidirectional communication would be slightly higher, since no ACK frame with following idles would be required. Beyond this, data transfer bandwidth scales directly with transmission clock speed, so that, for example, the data transfer rate over a double-speed link would be 100.369 * 2 = 200.738 MBps.

20.1.2. Fibre Channel Data Structures

The set of building blocks defined in FC-2 are:

Frame: A series of encoded transmission words, marked by Start of Frame and End of Frame delimiters, with Frame Header, Payload, and possibly an optional Header field, used for transferring Upper Level Protocol data.

Sequence: A unidirectional series of one or more Frames flowing from the Sequence Initiator to the Sequence Recipient.

Exchange: A series of one or more nonconcurrent Sequences flowing either unidirectionally from Exchange Originator to the Exchange Responder or bidirectionally, following transfer of Sequence Initiative between Exchange Originator and Responder.

Protocol: A set of Frames, which may be sent in one or more Exchanges, transmitted for a specific purpose, such as Fabric or N_Port Login, Aborting Exchanges or Sequences, or determining remote N_Port status.

Frames are the fundamental data transfer blocks in Fibre Channel; they contain a Frame header in a well-defined format and may contain a Frame payload. Frames are broadly categorized as either Data Frames, Link Control Frames (including Acknowledge (ACK) Frames), Link Response ("Busy" (P-BSY, F-BSY)), and "Reject" (P-RJT, F-RJT) Frames, indicating unsuccessful reception of a Frame, and Link Command Frames, including Link Credit Reset (LCR), used for resetting flow control credit values.

As stated above, each Frame is marked by Start of Frame and End of Frame delimiters. In addition to the transmission error detection capability provided by the 8B/10B code, error detection is provided by a 4-byte CRC value, which is calculated over the Frame Header, optional Header (if included), and payload. The %-byte Frame Header identifies a Frame uniquely and indicates the processing required for it. The Frame Header includes fields denoting the Frame's source N_Port_ID, destination N_Port_ID, Sequence_ID, Originator and Responder Exchange IDs, Frame count within the Sequence, and control bits. Every Frame must be part of a Sequence and an Exchange. Within a Sequence, the Frames are uniquely identified by a 2-byte counter field termed SEQ-CNT in the Frame Header. No two Frames in the same Sequence with the same SEQ-CNT value can be active at the same time, to ensure uniqueness.

When a Data Frame is transmitted, several different things can happen to it. It may be delivered intact to the destination, it may be delivered corrupted, it may arrive at a busy Port, or it may arrive at a Port that does not know how to handle it. Link Control Frames are used to indicate successful or unsuccessful reception of each Data Frame. The delivery status of the Frame will be returned to the source N_Port using Link Control Frames if possible. A Link Control Frame associated with a Data Frame is sent back to the Data Frame's source from the final Port that the Frame reaches, unless no response is required, or a transmission error prevents accurate knowledge of the Frame Header fields.

20.2. FIBER CHANNEL ROADMAP

A key aspect of the Fibre Channel architecture is a very active physical layer roadmap that follows the high growth rate of the storage industry. While Moore's Law says that the number of transistors per chip doubles every 18 months, the data capacity of disk drives has been doubling almost every year since the 1980s.[1] To keep up with this phenomenal growth rate, Fibre Channel has been doubling its line rate every two to three years. While Ethernet has traditionally increased its data rate by factors of 10, Fibre Channel takes an evolutionary approach that enables more affordable increases in speed.

The Fibre Channel Roadmap has three complementary speed roadmaps. The primary-speed-related technology is referred to as BASE-2 technology and doubles every few years, as seen in Fig. 20.6. Base-2 technology uses 8B/10B SERDES encoding and is the primary storage and server interconnect.[3] Base-10 technology started with 10GFC and uses 64B/66B SERDES encoding. The BaseT technology uses 4-dimensional Pulse Amplitude Modulation over 8 signal level (4D PAM-8) encoding to drive CAT cables. These three technologies create the foundation of Fibre Channel's physical layer.

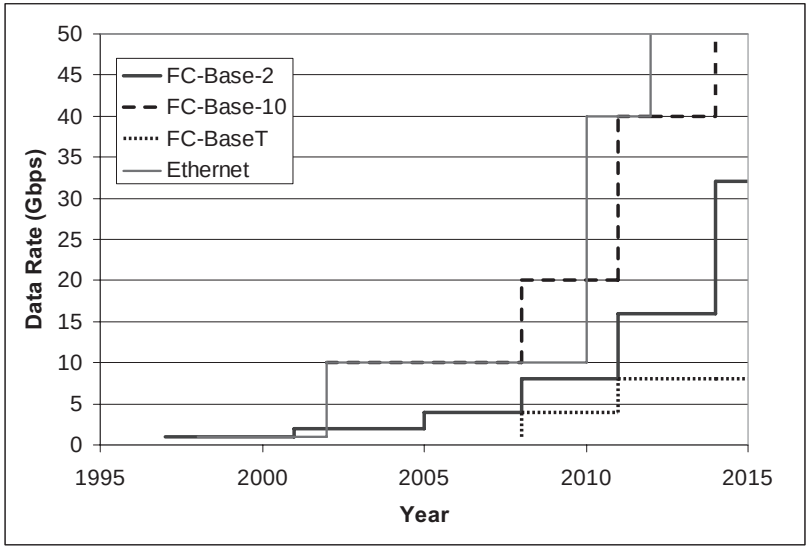


Figure 20.6 The three Fibre Channel speed technologies are shown when products have been or are expected to be deployed; Ethernet is shown for comparison.

[3]http://www.t11.org/t11/docreg.nsf/ufile/05-226v3, Fibre Channel—Physical Interfaces-2 (FC-PI-2), Greg McSorley.

Table 20.1
Fibre Channel Base-2 Speeds.

Product Naming	Throughput* (MBps)	Line Rate (GBaud)	T11 Specification Completed (Year)	Market Availability (Year)
1GFC	200	1.0625	1996	1997
2GFC	400	2.125	2000	2001
4GFC	800	4.25	2003	2005
8GFC	1,600	8.5	2007	2008
16GFC	3,200	17	2009	2011
32GFC	6,400	34	2012	Market Demand
64GFC	12,800	68	2016	Market Demand
128GFC	25,600	136	2020	Market Demand

*The throughput of the links includes data communications in both directions.

Table 20.2
Fibre Channel Base-10 Speeds.

Product Naming	Throughput* (MBps)	Line Rate (GBaud)	T11 Specification Completed (Year)	Market Availability (Year)
10GFC	2,400	10.52	2003	2004
20GFC	4,800	21.04	2007	2008
40GFC	9,600	42.08	TBD	Market Demand
80GFC	19,200	84.16	TBD	Market Demand
160GFC	38,400	168.32	TBD	Market Demand

*The throughput of the links includes data communications in both directions.

The BASE-2 speeds are used in over 99% of Fibre Channel ports shipped by 2007. BASE-2 products are the original Fibre Channel products and are always backward compatible with two generations of products. If a 4G port is plugged into a 1G port, the ports auto-negotiate to the highest available speed of 1GFC. Backwards compatibility is the main reason that BASE-2 ports will continue to dominate Fibre Channel deployments. Typically, the T11 Specification for the speed is released one to two years before products are released into the market. Finer details on the BASE-2 speeds are shown in Table 20.1.

The Base-10 speeds were created when Fibre Channel followed the lead of 10 Gigabit Ethernet that was defined by the IEEE. The 10G Fibre Channel standard shadowed the 10 Gigabit Ethernet movement and made minor changes to the standard so that the physical layer was nearly identical. Because of its high cost and incompatibility with Base-2 encoding, 10GFC has only been used for interswitch links since no end devices have adopted the Base-10 speeds. Fibre Channel plans to continue doubling the Base-10 speeds every few years as seen in Table 20.2.

Table 20.3
Fibre Channel Base-T Speeds.

Product Naming	Throughput* (MBps)	Line Rate (GBaud)	T11 Specification Completed (Year)	Market Availability (Year)
1GFC	200	1.0625	2006	Market Demand
2GFC	400	2.125	2006	Market Demand
4GFC	800	4.25	2006	Market Demand
8GFC	1600	8.5	TBD	Market Demand
16GFC	3200	17	TBD	Market Demand

*The throughput of the links includes data communications in both directions.

The third segment of the Fibre Channel roadmap is known as the Base-T roadmap and is based on copper cabling. FC-Base-T is designed to work over low-cost, copper, category (CAT) cables including CAT-5E, CAT-6 and CAT6a. Since FC-Base-T is the newest technology, 1/2/4G were immediately supported, as shown in Table 20.3. FC-Base-T links target cost-conscious customers that do not need to operate long-distance links. With no transceivers or fiber-optic patch-cords, FC-Base-T is designed for simple and easy installations.

The three branches of the Fibre Channel roadmap work in parallel to offer the most complete and low-cost solution. The majority of Fibre Channel links are Base-2 and have been optimized for low cost and adequate reach. The Base-10 links are most commonly used for ISLs, with typical data rates about 2.5 times higher than the Base-2 links. The FC-Base-T links were designed to increase adoption of Fibre Channel in the small to medium business (SMB) segment of the market, but no deployments of FC-Base-T were expected through 2008. Fibre Channel nomenclature for fiber-optic links is shown in Fig. 20.7.

The specification of links for a given fiber and speed depends on a number of specifications. Table 20.4 shows the link specifications from FC-PI-2 for Optical Multimode 2 (OM2) fiber that is called M5 fiber. Similar tables for OM1 or M6 fiber and OM3 or M5E fiber can be found in FC-PI-2. Table 20.5 shows the link specification for single-mode applications. The speed, operating distance, and rate tolerance define the capabilities of the link. Various transmitter and receiver specifications are also defined for a given link. Together, these specifications define the interoperability points and parameters of the link.

20.3. MULTIMODE LINK CONSIDERATIONS

The variety of speeds in Fibre Channel supports a variety of distances of multimode fiber. With three types of fiber being used at different speeds, 15 distances are supported in Fibre Channel as shown in Fig. 20.8. The supported distance for each link type is based on a variety of assumptions regarding the fiber,

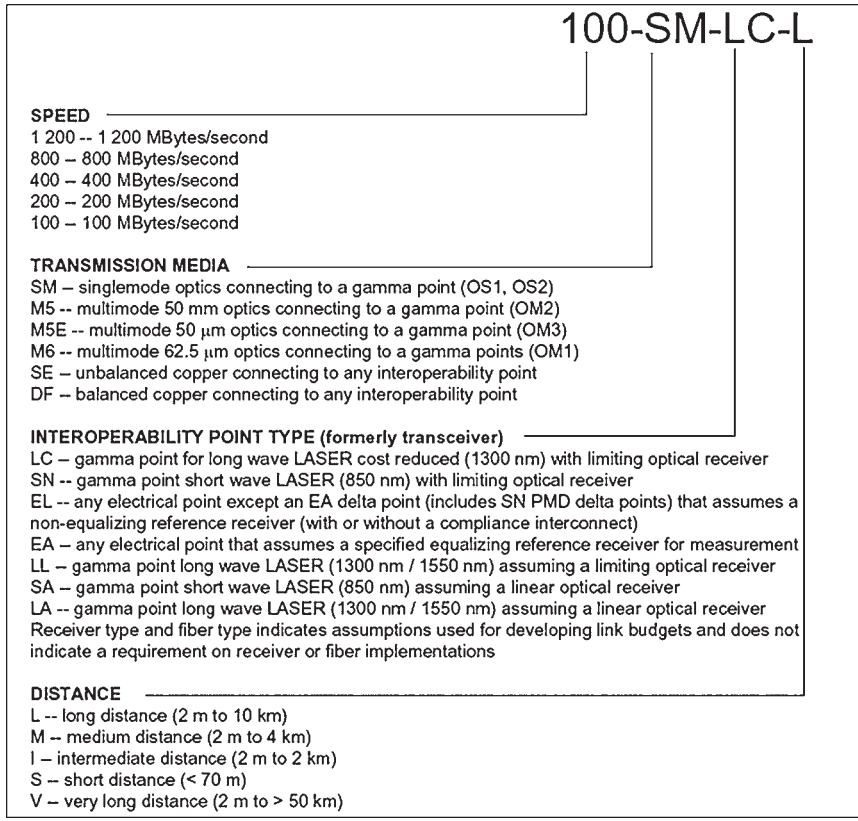

100-SM-LC-L
SPEED
1 200 -- 1 200 MBytes/second
800 – 800 MBytes/second
400 – 400 MBytes/second
200 – 200 MBytes/second
100 – 100 MBytes/second
TRANSMISSION MEDIA
SM – singlemode optics connecting to a gamma point (OS1, OS2)
M5 -- multimode 50 mm optics connecting to a gamma point (OM2)
M5E -- multimode 50 μm optics connecting to a gamma point (OM3)
M6 -- multimode 62.5 μm optics connecting to a gamma points (OM1)
SE – unbalanced copper connecting to any interoperability point
DF – balanced copper connecting to any interoperability point
INTEROPERABILITY POINT TYPE (formerly transceiver)
LC – gamma point for long wave LASER cost reduced (1300 nm) with limiting optical receiver
SN -- gamma point short wave LASER (850 nm) with limiting optical receiver
EL -- any electrical point except an EA delta point (includes SN PMD delta points) that assumes a non-equalizing reference receiver (with or without a compliance interconnect)
EA – any electrical point that assumes a specified equalizing reference receiver for measurement
LL – gamma point long wave LASER (1300 nm / 1550 nm) assuming a limiting optical receiver
SA – gamma point short wave LASER (850 nm) assuming a linear optical receiver
LA -- gamma point long wave LASER (1300 nm / 1550 nm) assuming a linear optical receiver
Receiver type and fiber type indicates assumptions used for developing link budgets and does not indicate a requirement on receiver or fiber implementations
DISTANCE
L -- long distance (2 m to 10 km)
M -- medium distance (2 m to 4 km)
I – intermediate distance (2 m to 2 km)
S -- short distance (< 70 m)
V – very long distance (2 m to > 50 km)

Figure 20.7 The nomenclature for Fibre Channel links.

transmitter, and receiver. Depending on the combination of factors for a given link, the supported link distance is a conservative estimate of how far the link should operate.

When 1GFC was defined as the first link based on VCSEL technology, the designers did not pay too much attention to the link length because the technology exceeded the needs for almost every application. T11 ended up defining 1GFC to support 300 meters on OM1 fiber. To keep the links cost effective, the bandwidth-length product (BWLP) (sometimes called the bandwidth-distance product) remained rather constant for each type of fiber as the speed increased to 8GFC as seen in Fig. 20.9. The BWLP of the link jumps considerably at the 10 Gbps speeds and when linear technology is used. The BWLP of the 10G links jumped

Table 20.4

This Table is from FC-PI-2 and Defines the Link Parameters for 1GFC, 2GFC and 4GFC on OM2 Fiber.

FC-0	Unit	100-M5-SN-I	200-M5-SN-I	400-M5-SN-I	Note
Sub clause		6.4	6.4	6.4	
Data rate	MB/s	100	200	400	
Nominal signaling rate	MBaud	1062.5	2125	4250	
Rate tolerance	ppm	±100	±100	±100	10
Operating distance	m	0,5–500	0,5–300	0,5–150	1
Fiber core diameter	µm	50	50	50	2
Transmitter (gamma-T)					
Type		Laser	Laser	Laser	
Spectral center wavelength, min.	nm	770	830	830	
Spectral center wavelength, max.	nm	860	860	860	
RMS spectral width, max.	nm	1,0	0,85	0,85	
Average launched power, max.	dBm				3
Average launched power, min.	dBm	–10	–10	–9	4
Optical modulation amplitude, min.	mW	0,156	0,196	0,247	5
Rise/Fall time (20%–80%), max.	ps	300	150	90	6
RIN_{12} (OMA), max.	dB/Hz	–116	–117	–118	7
Receiver (gamma-R)					
Average received power, max.	dBm	0	0	0	
Unstressed receiver sensitivity, OMA	mW	0,031	0,049	0,061	5,9
Return loss of receiver, min.	dB	12	12	12	
Rx jitter tolerance test, OMA	mW	0,064	0,107	0,154	
Stressed receiver sensitivity, OMA	mW	0,055	0,096	0,138	5,9,11
Stressed receiver vertical eye closure penalty	dB	0,96	1,26	1,67	9
Stressed receiver DCD component of DJ (at TX), min.	ps	80	40	20	
Receiver electrical 3 dB upper cutoff frequency, max.	GHz	1,5	2,5	5,0	8
Receiver eletrical 10 dB upper cutoff frequency, max.	GHz	3	6	12	8

Table 20.5

FC-0 Specifications for Single-Mode Links from FC-PI-2.

FC-0	Unit	100-SM-LC-L	200-SM-LC-L	400-SM-LC-L	400-SM-LC-M	800-SM-LC-L	800-SM-LC-I	Note
Data rate	MB/s	100	200	400	400	800	800	
Nominal signaling rate	MBaud	1,062,5	2,125	4,250	4,250	8,500	8,500	
Rate tolerance	ppm	±100	±100	±100	±100	±100	±100	10
Operating distance	m	2–10,000	2–10,000	2–10,000	2–4,000	2–10,000	2–1,400	
Fiber mode-field (core) diameter	µm							1
Transmitter (gamma-T)								
Type		Laser	Laser	Laser	Laser	Laser		
Spectral center wavelength, min.	nm					1,260	1,260	2
Side-mode suppression	dB	NA	NA	NA	NA	30	NA	
–20 dB spectral width	nm	NA	NA	NA	NA	1	NA	
Spectral center wavelength, max.	nm					1,360	1,360	2
RMS spectral width, max.	nm					NA		2
Average launched power, max.	dBm							3
Average launched power, min. Optical modulation amplitude, min.	dBm mW	–9,5	–11,7	–8,4	–11,2	–8,4 0,29	–8,4 0,29	4 2,5,13
Rise/Fall time (20%–80%), max.	ps	320	160	90	90	NA	50	6,12
RIN_{12} (OMA), max.	dB/Hz	–116	–117	–118	–120	–128	–128	7
Transmitter and dispersion penalty, max	dB	NA	NA	NA	NA	3.2	NA	14
Receiver (gamma-R)								
Average received power, max.	dBm	–3	–3	–1	–1	+0,5	+0,5	
Rx jitter tolerance test, OMA	mW	0,029	0,022	0,048	0,048	0,066	0,066	
Rx jitter tracking test, pk-pk amplitude	(kHz,UI)	NA	NA	NA	NA	(510,1) (100,5)	(510,1) (100,5)	15
Unstressed receiver sensitivity, OMA	mW	0,015	0,015	0,029	0,029	0,046	0,042	5,9,11
Return loss of receiver, min.	dB	12	12	12	12	12	12	
Receiver electrical 3 dB upper cutoff frequency, max	GHz	1,5	2,5	5,0	5,0	10	10	8

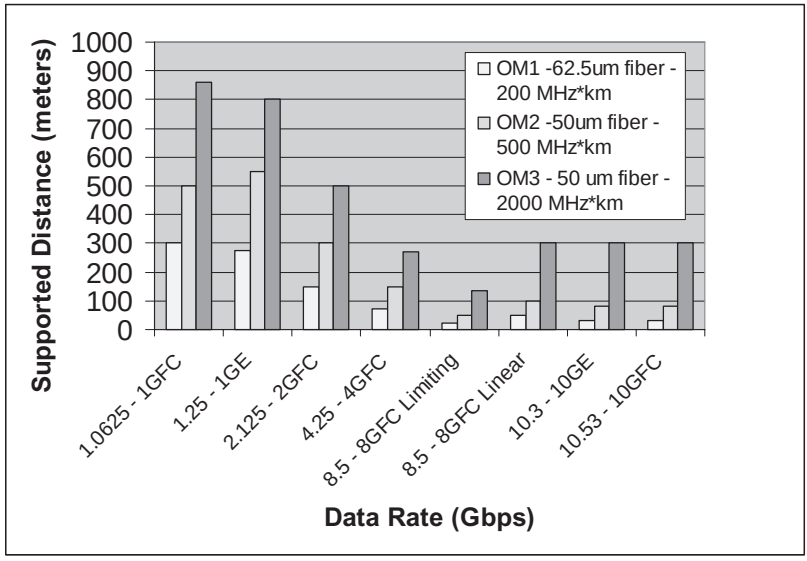


Figure 20.8 Supported distances of multimode links for Fibre Channel and Ethernet.

because of marketing requirements and not because of a technological breakthrough.

As increasing data rates have reduced the maximum link distance, new technologies have emerged which attempt to compensate for signal distortion. These include electronic dispersion compensation (EDC) and various linear and limiting receiver designs, as noted in Fig. 20.9. A more detailed description of these effects is provided in Chapter 7.

20.4. LINK POWER BUDGET ESTIMATION

With structured cabling becoming more common in large data centers, users may need to design links using a combination of patchcords and trunk cables. The trunk cables may be composed of OM3 fiber while the patchcords may be OM2 fiber. The supported distances in the previous section are meant to be over only one type of fiber. With the following power budget estimator, the user can calculate how far the link can be extended when different types of fibers are used in one link.

Supported distances are more complex when a link is composed of multiple patchcords with different types of fibers. To estimate the supported link length of fibers, a model has been devised that helps users determine the possible

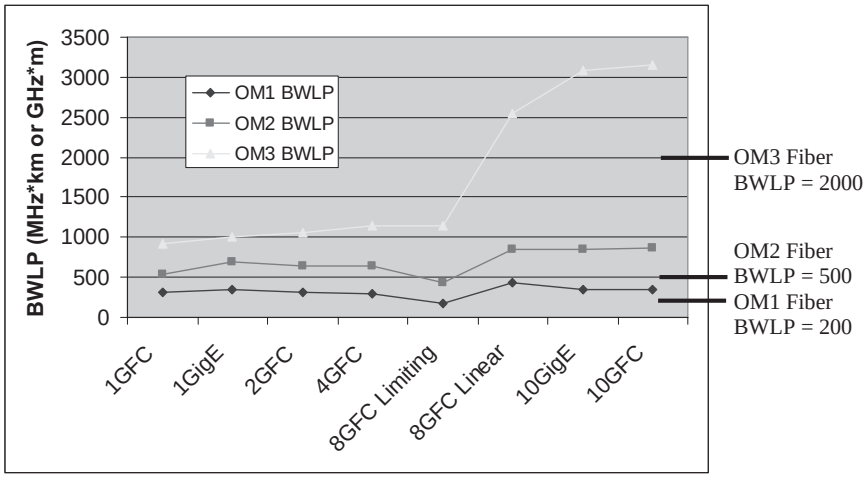


Figure 20.9 The bandwidth-length product for different types of optical links; note the sharp increase at 8GFC Linear and 10 Gbit/s data rates.

operating length of a multimode link. By dividing the link length by the power budget, the effective link attenuation is determined for a given speed and type of fiber shown in Table 20.6. The effective link attenuation is an approximation of all of the signal degradation that occurs on the link, including multimode dispersion, intersymbol interference, and 1.5 dB of connector loss. The effective link attenuation can help the user determine if a link is practical or if it can be extended.

For example, if the link is running at 400 MB/s and has 4 patchcords (10 meters of M5 (OM2) fiber, 60 meters of M5E (OM3) fiber, 35 meters of M5 fiber, and 6 meters of M5 fiber), the unused link power budget may be estimated as follows:

Link Power Budget – effective link attenuation = Unused link power budget

The Link Power Budget is 6.08 dB at 400 MB/s, and the link attenuation is determined by calculating the loss for the patchcords that comprise the link. A worksheet that shows the loss for each patchcord is presented in Table 20.7 and charted in Fig. 20.10.

The calculations show that less than half the link budget has been consumed in the first 111 meters of the link. If the user desired to extend the link, he or she could add new patchcords to the worksheet with the remaining link power budget.

Table 20.6
Effective Link Attenuation.

Link Type	Link Power Budget (dB)	Distance (meters)	Effective Link Attenuation (dB/km)
100-M6-SN-I	7	300	23.3
200-M6-SN-I	6	150	40.0
400-M6-SN-I	6.08	70	86.9
800-M6-SN-S	6	21	285.7
800-M6-SA-S	6.8	40	170.0
100-M5-SN-I	7	500	14.0
200-M5-SN-I	6	300	20.0
400-M5-SN-I	6.08	150	40.5
800-M5-SN-S	6	50	120.0
800-M5-SA-I	6.8	100	68.0
100-M5E-SN-I	7	860	8.1
200-M5E-SN-I	6	500	12.0
400-M5E-SN-I	6.08	380	16.0
800-M5E-SN-I	6	150	40.0
800-M5-SA-I	6.8	300	22.7

Table 20.7
Multimode Link Power Budget Example.

	Link Type	Distance (meters)	Effective Link Attenuation (dB/km)	Attenuation (dB)	Total (dB)	Distance (meters) 0	Unused Link Power Budget (dB) 6.08
Patchcord 1	400-M5-SN-I	10	40.50	0.41	0.41	10	5.68
Patchcord 2	400-M5E-SN-I	60	16.00	0.96	1.37	70	4.72
Patchcord 3	400-M5-SN-I	35	40.50	1.42	2.78	105	3.30
Patchcord 4	400-M5-SN-I	6	40.50	0.24	3.03	111	3.05

With 3.1 dB remaining, the link could be extended by 193 meters with M5E fiber, or 76 meters with M5 fiber.

This simple model assumes that the connector loss does not exceed 1.5 dB over the length of the link. Since patchcord connection losses are usually on the order of 0.25 dB, the link budget should be fine unless over 6 patchcord connections are used or some very lossy connections are in the link. One way to easily exceed the connection loss of 1.5 dB is to connect an OM1 fiber to an OM2 or OM3 fiber. With the core mismatch between the two fibers, losses of over 2 dB are expected, so users should not mix OM1 with OM2 or OM3 fiber. While this is not a formally

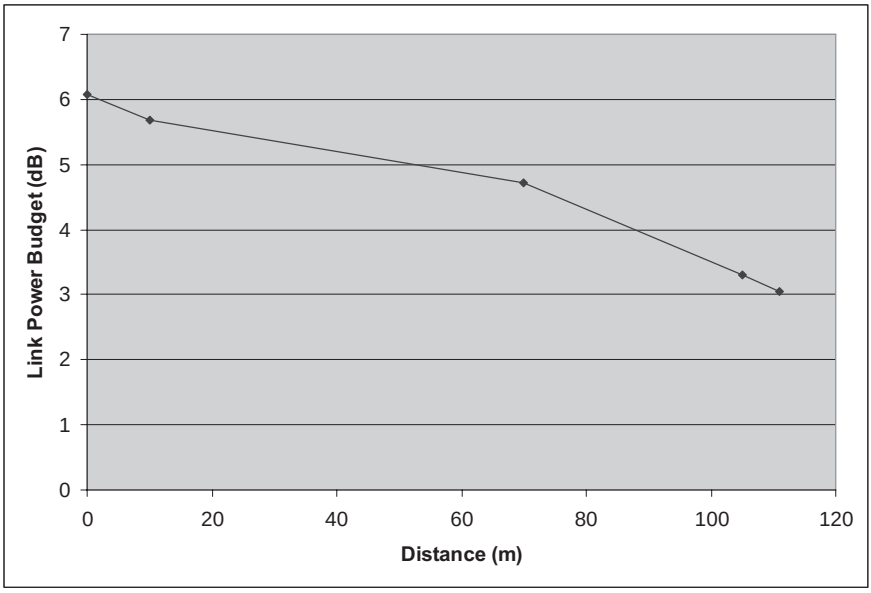


Figure 20.10 Example of link power budget vs. distance for different fiber types.

supported model, it is expected to cover a high percentage of links installed today because the links were defined conservatively in Fibre Channel and excess margin usually exists. Another model for calculating link length has been used in FC-PI-4. This model uses a graphical approach that results in the same link lengths; it can be found at http://www.t11.org/t11/docreg.nsf/ufile/07-155v3.

20.5 SINGLE-MODE LINK CONSIDERATIONS

While multimode links work well over distances of a few hundred meters, single-mode links are needed to span kilometers within cities or campuses. These links usually connect data centers or remote backup sites and are becoming more common as businesses plan for disaster recovery and business continuance. For highly available applications and services, companies often operate redundant single-mode links over previously dark fiber.

Two types of lasers have been standardized in Fibre Channel for single-mode applications: 1310-nm lasers and 1550-nm distributed feedback (DFB) lasers. The 1310 nm lasers are commonly Fabry-Perot (FP), but some 1310-nm VCSELs began shipping at these longer wavelengths in 2005. The 1310-nm lasers are limited by chromatic dispersion to about 10 km since they have a considerable spectral width. DFB lasers are very refined in the spectral domain and are thus

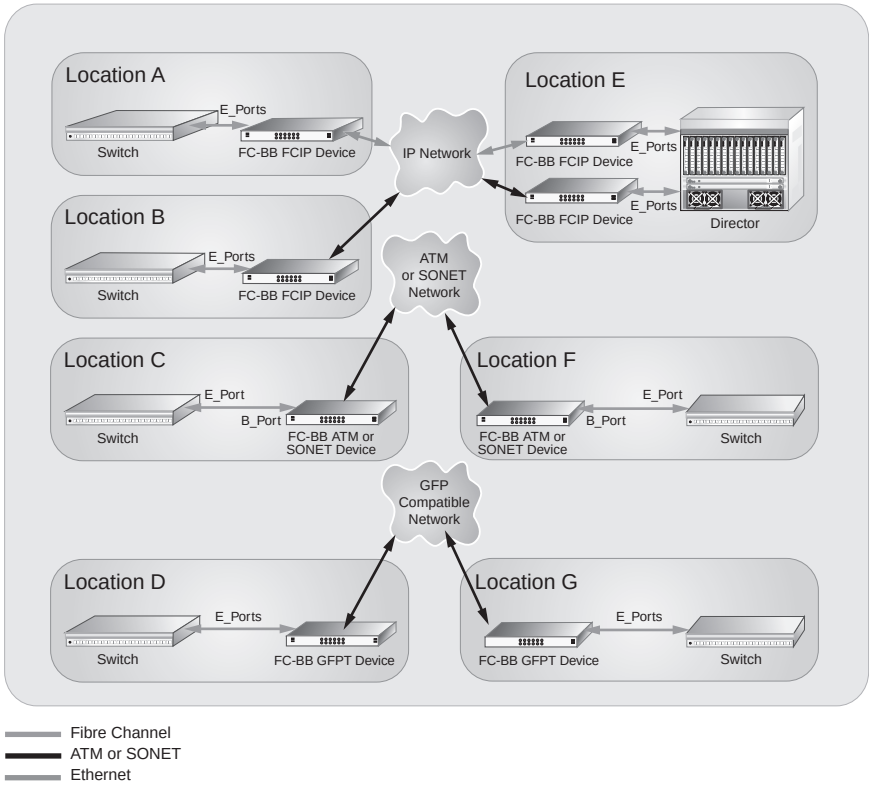


Figure 20.11 WAN interconnect devices.

limited by the optical power levels of the fiber after the links have gone beyond about 50 km. These links are capable of spanning most intracity distances.

The T11 Technical Committee standardized FP lasers as the first type of single-mode lasers supporting up to 10 km. As data rates increased, tighter and tighter restrictions on the spectral width of the FP lasers were required to maintain the 10-km link distance. At 4G, the spectral width of the laser had decreased to a little over 2 nm at the center wavelength of 1310 nm as seen in Figure 20.12. Since most FP lasers fail to meet this wavelength tolerance, vendors needed to use DFB lasers at considerably higher cost to reach 10 kms and maintain the 2 nm spectral width. To keep the 4G 10 km solution low cost, the standards body lowered the supported distance to 4 km and expanded the spectral widths of the link to 7 nm. This is illustrated by the so-called triple tradeoff curves as seen in Fig. 20.12; note that this general problem applies to many different link types, not just

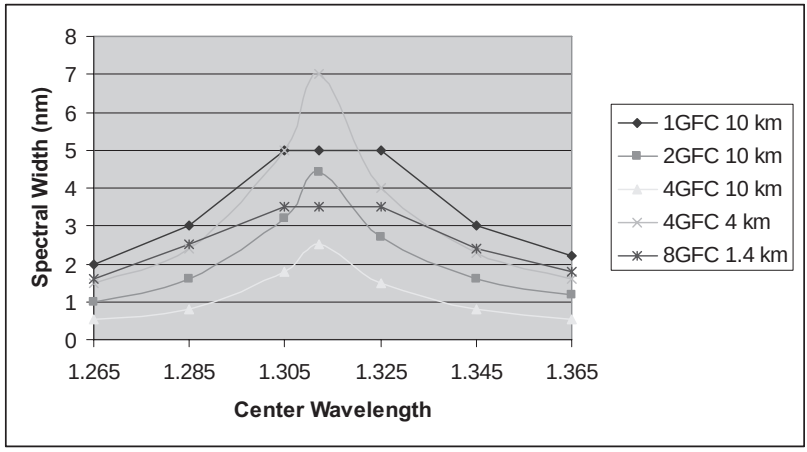


Figure 20.12 Triple trade off curves for Fibre Channel links.

Fibre Channel. At 8G, the link distance was further decreased to 1.4 km. This is an example of how the distance of the link (and thus the BWLP) was decreased at higher speeds to create a low-cost solution; whether the 1.4-km variant will be broadly adopted remains to be seen as of this writing.

Using a DFB laser source, it is possible to create links that span unrepeated distances of 50 km or more. In addition to having a tight spectral width of less than 0.1 nm, the DFB laser's well-controlled spot size can couple a large amount of power into a fiber. The maximum distances achieved with DFB lasers are usually limited by eye safety concerns; all Fibre Channel transceivers are Class 1 laser safe. Telecom transceivers that span hundreds of kilometers are not eye safe and have not been considered in Fibre Channel.

20.6. MAPPING TO UPPER LEVEL PROTOCOLS

The long distances discussed previously require optical fibers dedicated to Fibre Channel links. The high cost of installing or leasing dedicated fiber has meant that most applications over very long distances will employ some form of time and/or wavelength division multiplexing and may also encapsulate the data to operate over existing IP networks. On the other hand, it is also possible to encapsulate other types of data traffic in a Fibre Channel link. Before we address channel extension, we will first consider the issues related to link encapsulation.

The FC-4 level defines mappings of Fibre Channel constructs to ULPs. There are currently defined mappings to a number of significant channel, peripheral interface, and network protocols, including:

- SCSI (Small Computer Systems Interface)
- HIPPI (High Performance Parallel Interface)
- IP (the Internet Protocol) -IEEE 802.2 (TCP/IP) data
- SBCCS (Single Byte Command Code Set) or ESCON/SBCON/FICON

The general picture is of a mapping between messages in the ULP to be transported by the Fibre Channel levels. Each message is termed an Information Unit and is mapped as a Fibre Channel Sequence. The FC-4 mapping for each ULP describes what Information Category is used for each Information Unit, and how Information Unit Sequences are associated into Exchanges. The following sections give general overviews of the FC-4 ULP mapping over Fibre Channel for the IP, SCSI, and FICON protocols, which are three of the most important communication and I/O protocols for high-performance modem computers.

20.6.1. IP over Fibre Channel

Establishment of IP communications with a remote node over Fibre Channel is accomplished by establishing an Exchange. Each Exchange established for IP is unidirectional. If a pair of nodes wish to interchange IP packets, a separate Exchange must be established for each direction. This improves bidirectional performance, since Sequences are nonconcurrent under each Exchange, while IP allows concurrent bidirectional comunication. A set of IP packets to be transmitted is handled at the Fibre Channel level as a Sequence. The maximum transmission unit, or maximum IP packet size, is 65,280 **(x"FF00")** bytes, to allow an IP packet to fit in a 64-kbyte buffer with up to 255 bytes of overhead. IP traffic over Fibre Channel can use any of the classes of service, but in a networked environment, Class 2 most closely matches the characteristics expected by the IP protocol. The Exchange Error Policy used by default is "Abort, discard a single Sequence," so that on a Frame error, the Sequence is discarded with no retransmission, and subsequent Sequences are not affected. The IP and TCP levels will handle data retransmission, if required, transparent to the Fibre Channel levels, and will handle ordering of Sequences. Some implementations may specify that ordering and retransmission on error be handled at the Fibre Channel level by using different Abort Sequence Condition policies. An Address Resolution Protocol (ARP) server must be implemented to provide mapping between 4-byte IP addresses and 3-byte Fibre Channel address identifiers. Generally, this ARP server will be implemented at the Fabric level and will be addressed using the address identifier xFF FFFC.

20.6.2. SCSI over Fibre Channel

Fibre Channel acts as a data transport mechanism for transmitting control blocks and data blocks in the SCSI format. A Fibre Channel N_Port can operate

as a SCSI source or target, generating, accepting, and servicing SCSI commands received over the Fibre Channel link. The Fibre Channel Fabric topology scales better than the SCSI bus topology, since multiple operations can occur simultaneously. Most SCSI implementations in a storage device are over an Arbitrated Loop topology, for minimal cost in connecting multiple Ports. Each SCSI-3 operation is mapped over Fibre Channel as a bidirectional Exchange. A SCSI-3 operation requires several Sequences. A read command, for example, requires (1) a command from the source to the target, (2) possibly a message from the target to the source indicating that it is ready for the transfer, (3) a "data phase" set of dataflowing from the target to the source, and (4) a status Sequence, indicating the completion status of the command. Under Fibre Channel, each of these messages of the SCSI-3 operation is a Sequence of the bidirectional Exchange. Multiple disk drives or other SCSI targets or initiators can be handled behind a single N_Port through a mechanism called the Entity Address. The Entity Address allows commands, data, and responses to be routed to or from the correct SCSI target behind the N_Port. The SCSI operating environment is established through a procedure called Process Login, which determines operating environment such as usage of certain nonrequired parameters.

20.6.3. FICON

ESCON (Enterprise Systems Connection) has been the standard mechanism for attaching storage control units on IBM's zSeries eServer (previously known as S/390) mainframe systems since the early 1990s. ESCON channels were the first commercially significant storage networking infrastructure, allowing multiple host systems to access peripherals such as storage control units across long-distance, switched fabrics. In 1998, IBM introduced ESCON over Fibre Channel, termed FICON (Fibre Connection), which preserves the functionality of ESCON, but uses the higher performance and capability of Fibre Channel network technology. At the physical layer, FICON uses Fibre Channel. FICON also supports optical mode conditioners, which let single-mode transmitters operate with both single-mode and multimode fiber. This feature, which is also incorporated into Gigabit Ethernet, is not natively defined for Fibre Channel. FICON links at 1GFC, 2GFC, and 4GFC are currently available, with 8GFC links anticipated in 2008. This allows time-division multiplexing of up to 8 ESCON channels over a single FICON channel, a function once implemented as the FICON Bridge on some ESCON Directors.

At the protocol level, FICON is conceptually quite similar to SCSI over Fibre Channel, with a set of command and data Information Units transmitted as payloads of Fibre Channel Sequences. However, the FICON control blocks for the I/O requests, termed CCWs (Channel Control Words), are more complex and sophisticated than the SCSI command and data blocks. The FICON control blocks

Table 20.8
Fibre Channel Classes of Service.

Class 2	Duplicates the functions of a packet-switched network, allows multiple nodes to share links by multiplexing data as required
Class 3	Operates as Class 2 without acknowledgments

accommodate the different format and the higher throughput, reliability, and robustness requirements for data storage on these systems. The FICON physical layer also supports the use of mode conditioners at data rates up to 2.125 Gbit/s, in order to facilitate operating single-mode transceivers over multimode fiber (see Chapter 4).

Another difference between SCSI and FICON is that FICON currently does not support multi-hop or cascaded switch fabrics with more than two switches (similar to the ESCON protocol, which permitted only two switches, one of which was configured in static mode). In addition, FICON is optimized, in terms of overhead and link protocol, to support longer distance links, including links using DWDM (dense wavelength-division multiplexing), which allow transmission without performance droop out to 100 km. Longer distances may incur performance droop, although some specific applications can tolerate distances up to several hundred kilometers or even longer. Performance enhancements generally known as high-performance FICON implement various features such as buffer credit management, IU pacing, and modifications to storage control units. Combined with recent buffer credit enhancements on switches, this can significantly improve throughput on very long FICON links.

20.7. CLASS OF SERVICE

The Fibre Channel standard defines several classes of service; however, only two classes are typically used for transmitting different types of traffic under different delivery requirements. These are summarized in Table 20.8. Switches use connectionless routing and are characterized by the absence of dedicated connections. The connectionless Fabric multiplexes Frames at Frame boundaries between multiple source and destination N_Ports through their attached F_Ports. In a multiplexed environment, with contention of Frames for F_Port resources, flow control for connectionless routing is more complex than in the Dedicated Connection circuit-switched transmission. For this reason, flow control is handled at a finer granularity, with buffer-to-buffer flow control across each link. Also, a Fabric will typically implement internal buffering to temporarily store Frames that encounter exit Port contention until the congestion eases. Any flow control

Table 20.9
Fibre Channel Backbone.

Standard	Network	Mapping Name
FC-BB-1	ATM/SONET	FC-BB_ATM, FC-BB_SONET
FC-BB-2	Internet Protocol	FCIP
FC-BB-3	Generic Framing Procedure	FC-BB_GFPT
FC-BB-4	Pseudo Wire	FC-BB_PW
FC-BB-5	Fibre Channel Over Ethernet	FCoE

errors that cause overflow of the buffering mechanisms may cause loss of Frames. Loss of a Frame can clearly be extremely detrimental to data communications in some cases, and it will be avoided at the Fabric level if at all possible. In Class 2, the Fabric will notify the source N_Port with a BSY (busy) or a RJT (reject) indication if the Frame cannot be delivered, with a code explaining the reason. The source N_Port is not notified of nondelivery of a Class 3 Frame, since error recovery is handled at a higher level.

20.8. FIBRE CHANNEL OVER METROPOLITAN AND WIDE AREA NETWORKS

With telecom networks already running between corporate sites, T11 defined mappings of Fibre Channel onto multiple networks as seen in Table 20.9. Fibre Channel has proven to be very adaptable to metropolitan area networks (MANs) that span tens of kilometers and wide area networks (WANs) that span thousands of kilometers. Some generations of Fibre Channel switches have even integrated coarse WDM functions into the switch itself.

T11 mapped Fibre Channel to Asynchronous Transfer Mode (ATM) and Synchronous Optical Network (SONET) in FC-BB-1. The mapping of Fibre Channel onto the most popular telecommunication networks was mostly transparent to the Fibre Channel fabric. After initial protocol exchanges, the interswitch link acts identically to a long-distance fiber-optic link. The FC-BB_ATM or FC-BB_SONET device is the interface between the Fibre Channel network and the telecom network. The devices buffer data and control the flow between networks operating at different data rates.

FC-BB-2 took on a larger task of creating multiple virtual connections over IP networks. Fibre Channel over Transmission Control Protocol/Internet Protocol (FCIP) was a joint development between T11 and the Internet Engineering Task Force (IETF). T11 defined the means by which Fibre Channel networks interface with and connect across an IP network. The IETF defined the mapping and control required by TCP/IP in RFC 3821 and the FC frame encapsulation standard defined by RFC 3643. One of the main advantages of FCIP is that one physical link can

create multiple virtual connections to other end points. Figure 20.12 shows the differences between FCIP and FC-BB_ATM/SONET and FC-BB_GFPT.

FC-BB-3 mapped Fibre Channel onto Transparent Generic Framing Procedure (GFPT) networks. The FC-BB GFPT devices shown in Figure 20.11 are similar to FC-BB ATM devices in that the devices act as the interface between the two networks and are mainly transparent. In yet another protocol mapping, FC-BB-4 mapped Fibre Channel onto Pseudo-Wire (PW) networks. These protocols reverted back to being primarily transparent to the Fibre Channel network and look like a wire to the Fibre Channel devices that connect to them.

The latest mapping that is being developed as this chapter goes to press for FC-BB-5 is the Fibre Channel over Ethernet (FCoE) protocol. FCoE is designed to be a simple encapsulation protocol that encapsulates Fibre Channel frames that are sent over Ethernet networks. FC-BB-5 is intended to be used over lossless Ethernet networks that use flow control mechanisms at the physical layer. FCoE is designed as a low-overhead protocol in lossless networks in contrast to Small Computer Systems Interface over the Internet (iSCSI), which requires TCP processing in lossy networks. FCoE is basically attempting to use enhanced versions of Ethernet that will not drop frames and provide a reliable data network.

An important issue in extending Fibre Channel over distance is the use of credit-based flow control. Some protocols require multiple handshakes or data acknowledgments for the delivery of each frame (see, for example, the ESCON performance discussion in Chapter 14). Fibre Channel and FICON have reduced this overhead compared with ESCON; however, both still employ flow control mechanisms. Each end of the link has an allocation of buffer credits at the physical layer proportional to the size of the receive data buffer.

During link initialization, both ends of the link negotiate their maximum buffer size allocation. In order to avoid overflowing this buffer, whenever a Fibre Channel frame is transmitted, the far side of the link responds with an R_RDY command indicating there is sufficient receiver buffer space. As the link becomes very long, data frames and R_RDY commands may be stored on the link, and the end points must wait until these data structures complete a round trip on the link before transmitting additional data frames. The depletion of buffer credits, sometimes called buffer credit starvation, reduces the effective throughput of the link.

The maximum achievable distance before throughput degrades is proportional to the product of twice the link distance (allowing for a round-trip transport), the data rate, and the number of buffer credits. An example is shown in Fig. 20.13. Since long-distance applications typically require switches, the cost-effective design of ports with high buffer credits is important. Some switches employ buffer credit pooling, which allows them to reallocate unused buffer credits from short links to longer ones (this assumes that most of the attached links are short). Most commercial switches can allocate up to 2048 buffer credits per port and eliminate buffer credit starvation.

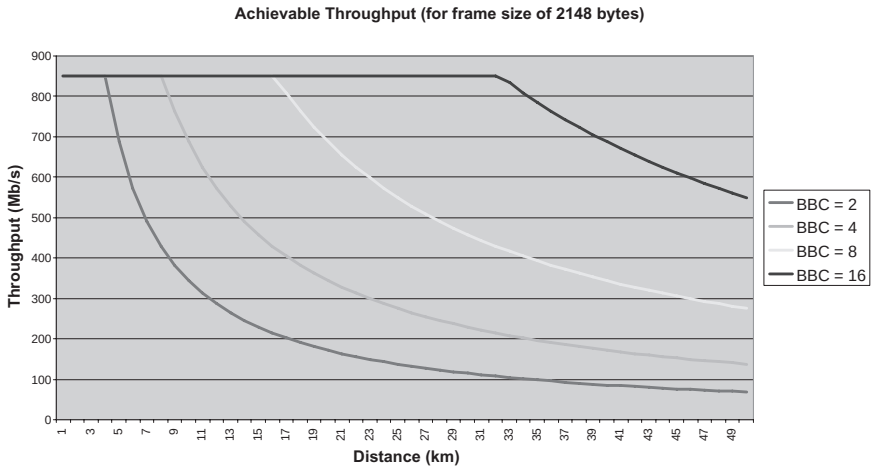


Figure 20.13 Example of performance droop due to credit-based flow control.

Other switches, channel extenders, or WDM equipment attempt to overcome this limitation by artificially generating R-RDY commands before frames reach the far end of the link. Known as "spoofing" the channel, this method requires additional link recovery mechanisms. Furthermore, there is an analogous credit-based flow control mechanism designed into the FC-4 layer; this must also be addressed in order to achieve high performance over long distances. Recent types of FICON channels have been designed to overcome these limitations.

20.9. CONCLUSION

This chapter has shown how Fibre Channel replaced the shared SCSI bus architecture with Gbps serial connections over long-distance fiber-optic connections. The serial connections enabled SANs that offered new functionality such as shared resources and virtualization. Fibre Channel connections in large data centers span multiple floors and permit data centers to be connected over tens or thousands of miles. The adaptable Fibre Channel protocol became the essential storage interconnect.

Fibre Channel has seen rapid evolution in speed and distance. From 1GFC in 1996 to 8GFC in 2008, Base-2 Fibre Channel has been the mainstay of Fibre Channel, and 16GFC is the next evolutionary step. Following 10GE, 10GFC has been used for ISLs that span distance and consolidate thousands of server and storage ports into a single Fibre Channel Fabric. To keep high-speed links low cost over long distances, Fibre Channel has defined linear technologies that extend the links with EDC. Single-mode fiber-optic links have been designed to

extend links to tens of miles, while mappings onto other networks have extended Fibre Channel over global distances. Fibre Channel has even standardized copper interfaces for low-cost solutions to meet the needs of virtually every customer.

This chapter has shown practical examples of transceivers and links. From the GBIC to the QSFP, Fibre Channel companies have helped define the most common datacom transceivers. The latest transceiver designed for 8GFC and 10GE applications is the SFP+[4]. SFP+ linear solutions are designed to extend the distance of these high-speed links while keeping costs low. For users designing links, the chapter showed how to calculate link length for structured cabling environments, even links when multiple fiber types were used. Fibre Channel has led the industry in many areas, from standardizing VCSEL solutions to defining low-cost, linear technology.

Fibre Channel has been designed for a specific task of providing the best interconnect for storage traffic. While some prophets have claimed that Fibre Channel is dead and that iSCSI will prevail, Fibre Channel continues to offer high value and reliable service. With Fibre Channel being a highly effective solution, users have no need to change to other technologies, and they keep Fibre Channel alive with their regular investments.

ADDITIONAL RESOURCES

The following web pages provide information on technology related to Fibre Channel, SANS and storage networking, and other high-performance data communication standards.

Hard copies of the standards documents may be obtained from Global Engineering Documents, an IHS Group Company, at http://global.ihs.com/. Electronic versions of most of the approved standards are also available from http://www.ansi.org and at the ANSI electronic standards store. Further information on ANSI standards and on both approved and draft international, regional, and foreign standards (ISO, IEC, BSI, JIS, etc.) can be obtained from the ANSI Customer Service Department. References under development can be obtained from INCITS (InterNational Committee for Information Technology Standards), at http://www.T11.org.

The following sources provide information on technology related to Fibre Channel, SANs, and storage networking.

http://webstore.ansi.org Web store of the American National Standards Institute. Soft copies of the Fibre Channel Standards documents.

http://global.ihs.com Global Engineering Documents, An IHS Group Company. Hard copies of the Fibre Channel standards documents.

---

[4]ftp://ftp.seagate.com/sff/SFF-8431.PDF, SFF-8431 Specification for Enhanced 8.5 and 10 Gigabit Small Form Factor Pluggable Module “SFP+”, Ali Ghiasi.

http://www.fibrechannel.org Fibre Channel Industry Association.
http://www.snia.org Storage Networking Industry Association.
http://www.storageperformance.org Storage Performance Council.
http://www.iol.unh.edu University of New Hampshire InterOperability Laboratory—Tutorials on many different high-performance networking standards.

REFERENCES

Benner, Alan. 2001. *Fibre Channel for SANs.* New York: McGraw-Hill C.

Clark, Tom. 1999. *Designing storage area networks: A practical reference for implementing Fibre Channel SAN*s. Reading, Mass.: Addison-Wesley Longman.

Decusatis, C. 1995. Data processing systems for optoelectronics. In *Optoelectronics for data communicaztion*, eds. R. Lasky, U. Osterberg, and D. Stigliani, 219–283. New York: Academic Press.

Farley, Marc. 2000. *Building storage networks,* New York: McGraw-Hill C.

Partridge, Craig. 1994. *Gigabit networking.* Reading, Mass.: Addison-Wesley.

Primmer, M. 1996, October. An introduction to Fibre Channel. *Hewlett-Packard J.* 47:94–98.

Tanenbaum, Andrew. 1989. *Computer networks.* Englewood Cliffs, N.J.: Prentice-Hall.

Widmer, A. X., and P. A. Franaszek. 1983. A DC balanced, partition block 8B/10B transmission code. *IBM J. Res. Dev.*: 27A40–451.

Case Study Design of Next Generation I/O for Mainframes

Courtesy of IBM Corporation

Application: Redesign the input/output (I/O) subsystem of a large enterprise server in response to changing workloads and processor performance.

Description: The original mainframe, or enterprise server, computer architecture was first established by the IBM System/360 in the 1960s. At the time, all of the server I/O was interconnected through massively parallel copper links, known as bus-and-tag connections. These links were limited to a maximum distance of 400 feet (122 m) by signal-to-noise ratio considerations, at a maximum data rate of 4.5 MByte/s. Reconfiguration was extremely difficult, especially since devices were commonly attached with dual links or "twin tailed" for redundancy. The copper cables were well over an inch in diameter and could not be bent around tight corners; combined with the distance restrictions, this meant that all peripheral devices had to be located in close proximity to the server, giving rise to the so-called glass house architecture in the data center.

While copper I/O was sufficient when typical system performance was on the order of tens of MIPS (millions of instructions per second), the available I/O bandwidth was quickly outpaced by processor growth. By the 1980s, as performance increased into the hundreds of MIPS, it was clear that a brute force approach of adding additional channels would not keep pace with bandwidth needs (especially given the upper limit of 256 I/O channels built into the server architecture). This led to the development of the first fiber-optic channels for the mainframe, known as ESCON (Enterprise Server Connection).

With a significant increase in data rate (up to about 17 MBtye/s accounting for system overhead) and unrepeated distances up to 3 km, ESCON provided the incremental bandwidth required to keep servers running near full utilization for several additional generations of processors. This was combined with the

introduction of a switched infrastructure and the multiple image facility (MIF), among the earliest channel virtualization systems for fiber optics. Subsequently, both server processing power and storage continued to grow, making further changes necessary in order to maintain a balanced system. One approach might have been to make more efficient use of the available bandwidth.

Consider a typical 4-Kbyte data block transfer on an ESCON channel at 17 MByte/s; this operation would require about 200 microseconds to transfer data and about 800 microseconds total to complete when we include the ESCON protocol overhead. It would be possible to complete the same data transfer using only 100 MByte/s of the channel capacity, and multiplex other workload over the remaining bandwidth using TDM or similar approaches. This results in only about a 20% improvement in transaction time, not enough to sustain more than perhaps one additional processor generation. A similar brute force approach would require adding more inexpensive, low-bandwidth ESCON channels to a single-server image; however, this does not scale well either. The infrastructure cost would increase with the addition of more I/O hardware, cables, patch panels, and switches, management complexity increases, and both system footprint and power consumption increase.

Analysis of these trends over time led to the requirement for another incremental step increase in I/O bandwidth, with the introduction of the 100-MByte/s FICON channels in the late 1990s. The new channel type meant that the server's 256-channel architecture could be preserved, while the increased bandwidth per channel meant that the server now had the equivalent of perhaps a hundred additional ESCON channels' worth of bandwidth at its disposal. For example, the initial release of FICON limited the server to 24 FICON channels, each of which could carry the equivalent of 8 ESCON channels at 50% channel utilization. This increased the effective number of ESCON channels per server from 256 to 360 channels. Raw numbers of channels was not the only benefit, however; the new channel architecture also needed to increase the channel start rate, from 500 I/O per second per channel to over 4000 I/O per second per channel. FICON also permitted the intermix of large and small data blocks on a channel, relieving some of the performance issues associated with small block transfers on ESCON. The number of unit addresses per channel was increased from 1 K to 16 K, and the unit addresses per storage control unit were also increased from 1 K to 4 K.

Subsequent releases have relieved the 24 FICON channel constraint, and modern mainframes now support considerably more than 256 channels through virtualization and other technologies. The FICON channel data rates have continued to scale, through the addition of 200 MB/s and 400 MB/s links, and will likely increase to 800 MB/s in the near future. However, the same principles are used today to calculate channel equivalency when a new channel structure is introduced.

Case Study Storage Area Network (SAN) Extension for Disaster Recovery

Courtesy of Ciena Corporation, in collaboration with Brocade

Application: Develop a disaster recovery solution to prevent lost or inaccessible data if the primary data center is lost; the end user is a $1 billion international business and technology consulting firm serving 43 states and provincial governments, many agencies of the U.S. federal government, and a number of Fortune 500 businesses.

Description: Disaster recovery solutions allow businesses to resume operation after they have experienced some natural or man-made disruption (such as software corruption, computer viruses, power failure, hurricanes, etc.). For a given application, it is necessary to determine factors such as the recovery time objective (RTO, how long can the system be unavailable), the recovery point objective (RPO, how much data loss is acceptable), and the network recovery objective (NRO, how long does it take to switch over the network). This allows the determination of a cost/recovery relationship so that the incremental benefit of spending additional disaster recovery resources can be determined.

With the proper network design, benefits such as resource sharing and virtualization enabled by a local storage area network (SAN) can be extended into the disaster recovery environment. There is an industry trend toward the interconnection and consolidation of local, independent SAN "islands," which had previously run autonomously. By forming SAN islands into a geographically distributed network, it is possible to achieve near real-time remote tape and disk mirroring using industry standard protocols. For many extended distance applications, asynchronous disaster recovery solutions of this type provide acceptable levels of RTO, RPO, and NRO.

In this particular case, the large amounts of data to be mirrored would have been prohibitive for a pure time-division multiplexed (TDM) environment such as SONET/SDH (OC-3 links operate at around 155 Mbit/s, while Fibre Channel

(FC) /FICON links operate at 1–4 Gbps). The requirement to interoperate with an existing FC/FICON environment, and the sensitivity of these protocols to transport delay over a public network, created further concerns with a SONET/SDH network, even considering potential use of GFP for FC/FICON encapsulation.

The solution involves using intelligent FCP/FICON directors (Brocade Silkworm) to interconnect SAN islands within a primary and secondary data center, including concatenation of lower data rate links over higher data rate interswitch links (ISLs) via TDM. The primary and remote sites were then interconnected over a fiber distance of around 55 km using a 32-wavelength metro WDM (Ciena Online metro). Dark fiber for local access networks is available in several high population metropolitan areas in the United States, such as New York, Chicago, Atlanta, Dallas, Denver, Los Angeles, Philadelphia, and Seattle.

The topology of these solutions parallels a conventional SONET/SDH network. There is a "core" FC/FICON network within a data center, and lower speed SAN traffic is concatenated at the network edge for transport across the "long haul" SAN extension (a dark fiber metro ring) over a WDM physical layer. Additional low-speed traffic concatenation can be done at the WDM equipment, which can optionally interpret the FC/FICON frame header to enforce quality of service and fault isolation; 50-ms protection switching is preserved on the dark fiber ring. When dark fiber is available at a reasonable cost, provisioning and commissioning of an extended SAN can be equivalent or faster than a SONET-based solution.

In order to ensure good performance, the switches must provide sufficient buffer credits on the extended distance interfaces. Some switches can be provisioned with buffer credit pooling (the ability to assign buffer credits to any switch port as required), while others require special high buffer count switch blades. Additional buffer credit management can sometimes be performed by the WDM equipment; the use of coarse WDM on switch blades has also been investigated for some products. To further reduce latency, some applications use "cut-through" switching (the switch does not store the entire frame; instead, frames are resent before the entire frame is received). If an error occurs in the frame, the switch sets the end-of-frame (EOF) delimiter to indicate that the frame is invalid.

21

Enterprise System Connection (ESCON) Fiber-Optic Link

Daniel J. Stigliani, Jr.
IBM Corporation, Poughkeepsie, New York

21.1. INTRODUCTION

The modern business computing environment, with its emphasis on dissemination of data in a client/server model, has placed tremendous demands on large enterprise servers such as the IBM eServer System z to improve not only data processing and server capability but also system interconnection capability. In the early 1990s, IBM introduced the first in a series of new large-scale servers that provided a new system structure and architecture (Enterprise Systems Architecture/390) for coupling multiple data processing systems together and Enterprise System Connection (ESCON) architecture to provide high-bandwidth interconnection capability for System/390 products and attachments. This was the beginning of the large-server interconnection network evolution into the modern information technology paradigm. This chapter provides an understanding of the ESCON interconnection from a system perspective and design consideration.

21.2. ESCON SYSTEM OVERVIEW

ESCON systems architecture is a total network interconnection system for large server complexes [1, 2]. ESCON encompasses fiber-optic technology links, serial data transfer, new link-level protocols, data encoding/decoding, new system transport architecture, and a new topology. The application for ESCON is intended as the backbone network that spans a customer's premises. In some cases it may be a machine room, whereas in other cases it could be a large multibuilding campus that may span 20 km or more.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

21.2.1. ESCON Topology

The topology chosen for ESCON is "switched point-to-point." It offers the highest throughput, excellent connectivity with minimal number of links, and the ability to grow the network in a nondisruptive manner. The switched point-to-point topology utilizes a central switch (director) to direct the network traffic to the various elements of the network [3] (note that these directors have been discontinued from IBM, although they remain available from other companies). The use of a director allows the connectivity of any unit on the network to any other unit on the network. The physical connections are point-to-point links that are ideally suited to fiber-optic technology. This topology enables the ability to isolate links in the network for failure analysis and repair. An n port nonblocking director can accommodate n/2 simultaneous conversations between end points in the switched point-to-point network.

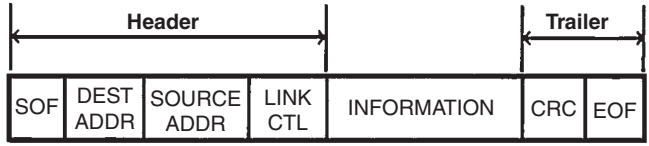
An important availability element of this configuration is that all servers and devices have two paths to each director. This configuration provides not only connectivity between servers and devices but also full redundancy and multipathing. For example, if any one of the links or directors becomes inoperative, there is an alternate path between the system and device. Also, by adding four links (two to each director), a new server can be included in the network nondisruptively, with immediate full connectivity.

21.2.2. ESCON Architecture and Channel

The IBM ESCON architecture establishes the rules and syntax used by the server to communicate to attached devices [4]. The architecture was defined to provide efficient transmission of data over long distances via a communication channel with a bit error rate (BER) of 10^{-10} (1 error in 10^{10} bits) or less. The architecture can be divided into two fundamental categories: device level and link level. The device level defines the rules for communication of a large server to an attached device using the facilities of the physical link. It defines data and control messages and the protocol to implement the server input/output(I/O) functions. The link-level architecture defines the actual transmission of information across the physical path. It defines the frame structure, type of frames, link initialization, exchange setup, data and control messages, address structure, and link error recovery.

21.2.2.1. Link Protocol

All information on the ESCON link is transferred within a frame structure or a sequence of special characters [4]. The ESCON frame is used to transport control and data information and is structured as shown in Fig. 21.1 [6]. The ESCON frame is delimited by a start-of-frame (SOF) and end-of-frame (EOF) ordered set



SOF:	Two character start-of-frame delimiter.
DEST ADDR:	Two byte destination address of frame.
SOURCE ADDR:	Two byte source address of frame.
LINK CTL:	One byte of link control information.
INFORMATION:	Zero to 1028 bytes of data.
CRC:	Two byte cyclic-redundancy-check information.
EOF:	Three character end-of frame delimiter.

Figure 21.1 ESCON frame structure.

of characters, respectively. The SOF and EOF are unique sets that are also used by the director to establish a connection, continue a connection, or disconnect after completion of the frame transmission. The SOF delimiter is composed of two characters (20 bits). The next 16 bits (before encoding) are reserved for the destination address, the next 16 bits (before encoding) contain the source address, and the next 8 bits (before encoding) are a link control field. The link control field indicates the type and format of the frame. The four fields above are known, as a group, as the link header.

The next field following the header (Fig. 21.1) is the information field, which may contain data or system information and can vary from 0 to 1028 bytes. The link trailer consists of two fields, cyclic-redundancy-check field (CRC), and the EOF field. In order to ensure the data are received correctly, a CRC is generated at the transmitter and included in the frame as a 16-bit CRC field. The receiving device uses the CRC to verify the information field. The use of fiber-optic technology has ensured that link errors from external stimulus are extremely low and that the random bit error rate of the optical link due to receiver noise is less than 10^{-15}. Based on these low error rates, the recovery approach is to retransmit the frame if an error has occurred. Because this happens so seldom, the system performance is not affected by this recovery approach. The EOF field is a three-character (30-bit) field that signifies the end of the frame. The data between the SOF and EOF delimiters are modulo of 8 bits before transmission and encoded into 10-bit characters for transmission on the link.

The architecture also defines an ordered set of sequences that can be transmitted over the link in the presence of a very high error rate condition (in which frames cannot be transmitted correctly). Each sequence contains a continuous

repetition of an ordered set to maximize the likelihood that a sequence will be correctly recognized. Some typical sequences are not operational sequence, in which a link-level facility (at the server or device) cannot interpret a received signal, or offline sequence, in which the appropriate link-level facility is indicating that it is offline with respect to sending any information. These and other sequences are interpreted at a level above the link layer, and appropriate action is taken by the server.

An idle character is always sent on the link when no frames or control sequences are being sent. The idle character is a special ordered set of bits (named K28.5) [7]. Also, idles are sent between frames as well. The idle sequence ensures that the receiver is both in bit and character synchronization with the transmitter. If the receiver becomes out of synchronization with the transmitter, the architecture has defined a set of rules and procedures whereby synchronization can be reacquired [7].

21.2.2.2. Data Encoding/Decoding

High-speed fiber-optic receivers perform best over environmental and manufacturing variation when they are AC coupled. The ESCON optical receiver is designed in this manner. In order to prevent DC baseline wander, it is important to ensure that the information on the link is encoded from the normal nonreturn to zero (NRZ) computer code to a DC-balanced code. Several codes (e.g., Manchester and 4B/5B) were investigated, and an 8B/10B code was chosen for ESCON. This technique was chosen because it provides the most robust code and a minimum bandwidth overhead (25%). For example, the 8B/10B code contains special control characters that will not degrade into a another valid character with single-bit errors.

The 8B/10B encoding transforms a byte (8 bit) of information at a time into a 10-bit transmission character. The 10-bit character is sent serially bit by bit over the fiber-optic link and decoded at the receiver into the original 8-bit byte. Conceptually, the 256-bit combinations of the 8-bit byte are mapped into a subset of the 1024 10-bit characters such that the maximum run length of 1s or 0s is 5. Special control characters and sequences (e.g., idle, SOF, and EOF) are defined that are not derived from the 8-bit original but are meaningful only as architected control and definition characters. For example, the +K28.5 idle character (0011111010) is unique, and there is no valid data character with this 10-bit sequence [7]. A single-bit error will not result in a valid 10-bit character. Only 536 of the 1024 possible characters are valid. All others will cause an architected error condition.

The running disparity (difference between the number of 1s and 0s in a character) is continually monitored to ensure a DC balance. If disparity exceeds the

bounds, an error condition occurs. The 8B/10B code is well behaved with regard to DC balance, and the number of transmissions between 0s and 1s is sufficient to ensure that the receiver, retiming, and character recognition circuits can reliably perform the required functions.

21.2.2.3. Bit Error Rate Thresholding

The architecture is tolerant of bit errors on the link that may be detected as code violation, sequencing, or CRC errors [8]. A code violation occurs when an invalid transmission character is received. A sequencing error occurs when a sufficient number of consecutive special ordered sets (discussed earlier) cannot be transmitted without error. Finally, a CRC error occurs when the CRC result of the received frame contents is not equal to the expected value.

For the link design, the number of retries due to link errors has a negligible effect on link performance. However, as the rate of retries increases beyond a threshold value the degradation of the link may be noticeable. A report is generated when the specified threshold is reached on a link for further analysis and maintenance. The threshold for ESCON is set at 1 error in 10^{10} bits. At this level the link performance is still tolerable, and maintenance can be deferred until a convenient time. Beyond this level, the server will begin to realize degraded performance on that link.

The actual measurement is done by counting the number of code violation events within a specified time. A bit error will likely cause more than one code violation. Consequently, the concept of an error burst has been developed. To prevent a single-bit error from causing multiple error counts, one or multiple code violations within a 1.5-second period are considered as one error burst for the threshold count. Fifteen or more error bursts within a 5-min period will result in a threshold error recorded by the server. The threshold count is reset when the threshold is reached, or every 5 min, whichever occurs first. Detailed information is given in Ref. [8].

21.3. ESCON LINK DESIGN

The transition of computer interconnection from parallel copper technology to a radically different technology (“serial” fiber optics) generated many questions and concerns. Most of the concerns centered around the reliability of the link in a computer data center environment. Can the technology meet the stringent reliability requirement for both bit errors on the link and hardware failures? The fiber-optic link must perform equal to or better than the copper links it replaces. The ESCON link design [9] and component selection were made to achieve both high data rate and reliability.

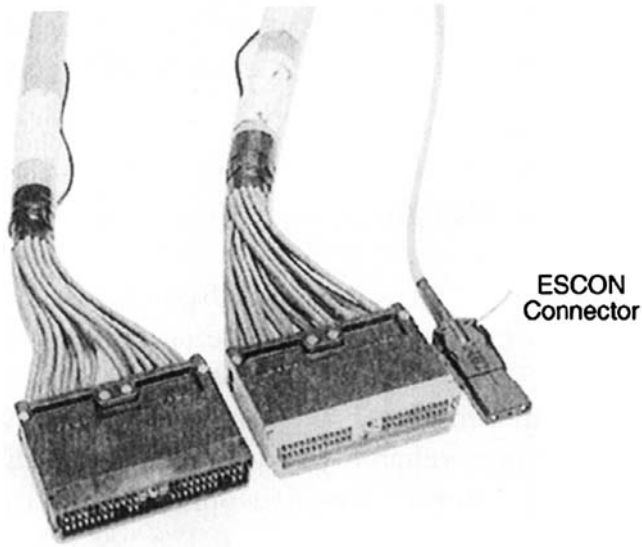


Figure 21.2 Parallel copper and ESCON channel cables [9]. (Copyright 1992 by International Business Machines Corporation, reprinted with permission.)

21.3.1. Multimode Design Considerations

The multimode ESCON link replaced a parallel (8-bit wide) copper coaxial cable link that had proven reliability and performance. Any replacement of the copper link must be easier to use, offer higher data rate and distance performance, be smaller in size, lighter in weight, and equal or better in reliability. Figure 21.2 depicts the size reduction of the interconnection cable and connector of ESCON compared to the equivalent (two) parallel copper cables and connectors it replaces.

The optical link must extend throughout a campus environment (typically 2 or 3 km) and achieve very reliable data transfer. A optical link BER design of 10^{-15} for the worst-case (longest length) link was chosen.

21.3.1.1. Major Components

The major components of the optical link are illustrated in Fig. 21.3. The serializer (typically implemented in Complementary metal oxide semiconductor [CMOS] technology) takes the 10 parallel bits of 8B/10B encoded data and serializes the data into a 200-Mb/s rate serial bit stream, whereas the deserializer performs the complementary function. The deserializer also includes the retiming function, which extracts the clock from the serial data. The derived clock is used

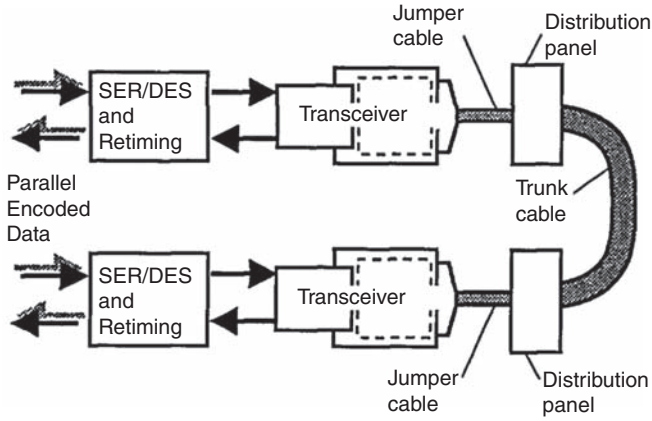


Figure 21.3 Block diagram of fiber-optic link elements [9]. (Copyright 1992 by International Business Machines Corporation, reprinted with permission.)

to latch and reshape the serial data prior to deserialization. The transmitter uses a light-emitting diode (LED) operating at 1300 nm, and the receiver uses a positive-intrinsic-negative (PIN) photodiode. Both devices are made of InGaAsP quaternary material. The 1300-nm LED was chosen because this wavelength is at the optimum attenuation and bandwidth of multimode fiber and has excellent reliability and low cost.

The jumper cable is a two-fiber (one inbound and one outbound), rugged, yet flexible, cable assembly that uses an aramid fiber strength member. The ESCON connector is a low-profile, polarized, push-on connector that latches into a transmitter receiver subassembly (TRS) or coupler assembly. The fiber used in the jumper is multimode 62.5/125 μm, and the ferrules are made of zirconia ceramic material. The ESCON link is designed to be used with either 62.5/125 or 50/125 μm multimode trunk fiber. The use of 62.5/125 μm trunk supports a link length of 3 km, whereas the 50/125 μm trunk fiber supports a 2-km link distance. The difference in distance capability is due to the additional loss associated with connecting a 62.5/125 μm jumper fiber to a 50/125 μm trunk fiber.

21.3.2. Single-Mode Design Considerations

The new long-distance ESCON link, called ESCON XDF, uses a long-wavelength laser as the source and single-mode optical fiber (SMF). The single-mode fiber chosen is the same as that used by the telecommunications industry and generally available. This is an important consideration because these long distances typically will traverse right of ways and likely the fiber is owned by another company (e.g., posts, telephone, and telegraphs; local telephone provider;

and power company). In general, the computer customer is not interested in fiber optics as an entity but only as a means of efficient communication within his or her network. To ensure ease of use and not require of the customer anything more than the base link requirements, the XDF must be an international class 1 laser safety product. This category allows unrestricted access by uncertified laser personnel because the product conforms with "eye safe" government and industry criteria.

The jumper cables use 9/125-μm fiber, whereas the trunk can use either 9- or 10-μm core fiber. There is no distance penalty associated with the use of 10-μm core trunk fibers.

The XDF feature provides a 20-km link capability at 200 Mb/s without the use of repeaters. The link distance is a function of the optical loss budget and is a tradeoff of laser transceiver cost and complexity versus distance. The laser power output is maintained at a low enough power level to ensure compliance with Class 1 laser safety standards.

The laser transceiver discussed in this chapter is a second-generation transceiver that utilizes the single-mode asynchronous transfer mode industry standard module package with the FCS connector. The prior version was a single-mode ESCON connectorized module that is no longer in production. It was designed and produced by IBM because no industry product at that time could meet the requirements of System/390 servers. The new and original laser transceivers are fully compatible and have similar specifications.

21.3.3. Multimode Link Design and Specification

The ESCON link budget elements are grouped into two major categories:

1. Cable plant
   The cable plant loss includes connector loss, fiber attenuation, higher order mode loss, and splices.
2. Available power
   The available power is the resultant optical power available for the link after the optical budget associated with the transmitter and receiver is adjusted for link losses such as
   - Fiber dispersion penalty (modal and spectral)
   - Retiming penalty
   - BER specification conversion from 10^{-12} to 10^{-15}
   - LED end-of-life degradation
   - Transceiver coupling variation
   - Data dependency

Link parameters are defined into these categories to allow maximum flexibility over the elements that can be controlled by the user (e.g., fiber attenuation) and

incorporate into the available power those elements that are difficult or cannot be controlled (e.g., fiber dispersion) by the user.

The elements of the available power budget are statistically summed to yield a resultant available power as a distribution with a mean and standard deviation. The following condition must be satisfied for the link to meet its design criteria as follows,

$$U_{av} - n\sigma_{av} \geq C_t, \tag{21.1}$$

where U_{av} is the available power, n is the number of standard deviations, σ_{av} is the standard deviation of the available power, and C_t is the total cable plant optical loss. For an ESCON link $n = 3$ (3 σ design) for the longest link allowed in the configuration at a BER of 10^{-15}. The resultant mean and standard deviation for the available power is determined using a Monte Carlo technique to sum the various elements. This was done because all the parameter distributions are not necessarily Gaussian, and in fact the transmitter output power and receiver sensitivity are truncated distributions. The use of a 3-σ design point for the worst-case link (3 km for 62.5-μm trunk and 2 km for 50-μm trunk) ensures that all shorter links are designed conservatively and the risk of an install link budget failure is extremely remote.

Table 21.1 illustrates the resultant specification of the cable plant to ensure the multimode link operates in accordance with the link design requirements. The maximum link loss is established at 8 dB independent of link configuration. The loss budget was maintained at 8 dB by adjusting the fiber bandwidth and in turn the dispersion penalty. The standard 2-km 62/125 μm link uses 500 MHz-km fiber, whereas the 2-km 50/125 μm and 3-km 62.5/125 μm link use a higher bandwidth (800 MHz/km) grade of fiber. This allows the customer maximum flexibility to adjust his or her configuration to the environment. The user can trade off number and connector quality with fiber attenuation and length to achieve an optimized installation.

Table 21.1

ESCON Maximum Link Loss (at 1300-nm wavelength).

Maximum Link Length (km)	Maximum Link Loss (dB)	Truck Fiber Core Size (μm)	Minimum Truck Modal Bandwidth (MHz/km)
2.0	8.0	62.5	500
2.0	8.0	50.0	800
2.0–3.0	8.0	62.5	800

Note: From Ref. [11]. The maximum link length includes both jumper and truck cables. The maximum total jumper cable length cannot exceed 244 m when using either 50/125 μm truck fiber or when a 62.5/125 μm link exceeds 2 km.

Table 21.2
ESCON XDF Maximum Link Loss (at 1300-nm wavelength).

Maximum Link Length (km)	Maximum Link Loss (dB)	Truck Fiber Core Size (μm)
20.0	14.0	9–10

Note: From Ref. [11]. The maximum link length includes both jumper and truck cables. The maximum of a single-mode jumper cable is 4 m. In a single-mode truck cable, distance between connectors or splices must be sufficient to ensure that only the lowest order bound mode propagates.
Single-mode connectors and splices must meet a minimum return loss specification or 28 dB. The minimum return loss of a single-mode link must be 13.7 dB.

21.3.4. Single-Mode Link Design and Specification

The single-mode link design follows the same approach used for the multimode design. The jumper fiber is 9/125 μm. The XDF link supports both 9- or 10-μm core fiber without any effect on distance. The excess loss (approximately 0.2 dB) associated with the coupling of a 9-μm core jumper to a 10-μm core trunk fiber is included in the available power category and is transparent to the overall link budget. The dispersion penalty of the fiber due to spectral width of the laser is small and has also been accounted for in the available power budget along with any effects due to laser mode hopping and relative intensity noise [10]. All these time domain effects are relatively small for a 200-Mb/s single-mode link and are included as a 1.5-dB fixed (no distribution) "AC optical path" penalty. The single-mode link specification is given in Table 21.2. A maximum link length of 20 km can be achieved with a maximum optical cable plant loss budget of 14 dB for the cable plant.

In order to ensure that the laser is well behaved under all operating conditions, it is important to minimize any optical reflections occurring in the cable plant. This is done by specifying that all connections and splices in the link have a minimum return loss of 28 dB. Mode partition noise in the XDF link is alleviated by specifying that no jumper less than 4 m may be used. The minimum length in conjunction with the specified cutoff wavelength of the fiber ensures that only the lowest order bound mode propagates in the jumper. Likewise, the trunk installer must ensure that any connectors or splices in the trunk meet the return loss specification and that all connectors or splices are placed sufficiently apart so that only the lowest order mode is propagating prior to any connectors, splices, or other optical discontinuities.

21.3.5. Multimode Optical Output Interface

The optical coupled light specifications required for an ESCON link are given in Table 21.3. The parameters specified will allow the maximum distance require-

Table 21.3
Multimode Optical Output Interface Specifications.

Parameter	Minimum	Maximum	Unit
Average power[a,b]	–20.5	–15.0	dBm
Center wavelength	1280	1380	nm
Spectral width (FWHM)		175.0	nm
Rise time (t_r) (20–80%)[a,c]		1.7	ns
Fall time (t_f) (80–20%)[a,c]		1.7	ns
Eye window[a]	3.4		ns
Optical output jitter[d]		0.8	ns
Extinction ratio[a,e]	8		dB
t_r, t_f at optical path output[c,f]		2.8	ns

Note: From Fef. [11].
[a]Based on any valid 8B/10B code. The length of jumper cable between the output interface and the instrumentation is 3 m.
[b]The output power shall be greater thatn – dBm through a worst-case link as specified in Table 21.1. Higher order mode loss (HOML) is the difference in link loss measured using the device transmitter compared to the loss measured with a source conditioned to achieve an equilibrium mode distribution in the fiber. The transmitter shall compensate for any excess HOML occurring in the link (e.g., HOML in excess of 1 dB for a 62.5-μm link).
[c]The minimum frequency response bandwidth range of the optical waveform waveform detector shall be 100 kHz to 1 GHz.
[d]The optical output jitter includes both deterministic and random jitter. It is defined as the peak-to-peak time-histogram oscilloscope value (minimum of 3000 samples) using a 2^7-1 pseudo-random pattern or worst-case 8B/10B code pattern. The transmitter output light is coupled to a PIN photodiode O/E converter (e.g., Tektronix P6703A or equivalent) via a 3-m cable and jitter measured with a digital sampling oscilloscope [13].
[e]Measurement shall be made with a DC-coupled optical waveform detector that has a minimum bandwidth of 600 MHz and whose gain flatness and linearity over the range of optical power being measured provide an accurate measurement of the high and low optical power levels.
[f]The maximum rise or fall time (from, e.g., chromatic, modal dispersion, etc.) at the output of a worst-case link as specified in Table 21.1. The 0 and 100% levels are set where the optical signal has at least 10 ns to settle. The spectral width of the transmitter shall be controlled to meet this specification.

ments and loss budget, as specified in Table 21.1 with a BER of 10^{-15}. The light source is an incoherent light-emitting diode.

21.3.6. Multimode Input Optical Interface

The input optical interface specifications are given in Table 21.4. A loss-of-light function and operation is specified for link failure indication and diagnostic use. The design of the machine receiving this information determines how this state change information is utilized.

Table 21.4
Multimode Optical Input Interface Specifications.

Parameter	Minimum	Maximum	Unit
Sensitivity[a,b]		–29.0	dBm
Saturation level[a]	–14.0		dBm
Acquisition time[c]		100	ns
LOL threshold[d]	–45	–36	dB
LOL hysteresis[d,e]	0.5		dB
Reaction time for LOL state change	3	500	μs

Note: From Ref. [11].
[a]Based on any valid 8B/10B code pattern measured at, or extrapolated, 10^{-15} BER measured at center of eye. This specification shall be met with worst-case conditions as specified in Table 14.3 for the output interface and Table 21.1 for the fiber-optic link. This value allows for a 0.5-dB retiming penalty.
[b]A minimum receiver output eye opening of 1.4 ns at 10^{-12} should be achieved with a penalty not exceeding 1 dB.
[c]The acquisition time is the time to reach synchronization after the removal of the condition that caused the loss of synchronization. The pattern sent for synchronization is either the idle character of an alternating sequence of idle and data characters.
[d]In direction of decreasing power: If power > –36 dBm, LOL state is inactive; if power < –45 dBm, LOL state is active. In direction of increasing power: If power < –44.5 dBm, LOL state is active; if power > –35.5 dBm, LOL state is inactive.
[e]Required to avoid random transitions between LOL being active and inactive when input power is near threshold level.

21.3.7. Multimode Fiber-Optic Cable Specification

The two optical fibers are assembled into a duplex optical cable assembly for the jumper and assembled into pairs for the trunk. The jumper cable assembly is terminated in the ESCON duplex fiber-optic connector. The trunk cable, however, is usually installed in high-count configurations (e.g., 12, 24, 36, 72, and 144 fiber counts) by professionals skilled in the art of fiber-optic installation. The planning and installation of the trunk is reviewed in Section 21.5. The two fibers in a jumper cable are assembled as illustrated in the cable cross section (Fig. 21.4). The cable assembly is nonmetallic and uses aramid fiber as the strength member. All the elements are encased in a flexible polyvinyl chloride (PVC) jacket.

The optical specifications in this section are associated primarily with the fiber and are necessary to ensure that the link meets its performance objectives. They also ensure consistency among various ESCON-compatible devices.

21.3.7.1. Multimode Jumper Cable Assembly

The MMF jumper cable is only offered in a 62.5/125 μm fiber configuration, and the optical specifications are given in Table 21.5. The cable jacket color is

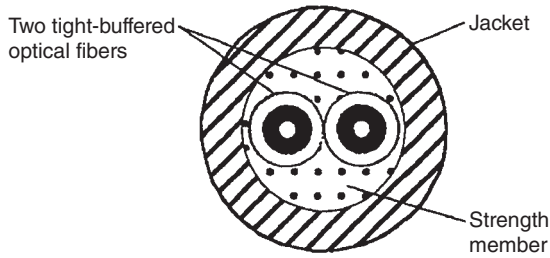


Figure 21.4 Multimode jumper cable construction [12].

Table 21.5
Multimode (62.5/125 μm) Jumper Cable Specifications.

Parameter	Specification
Fiber type	Graded index with glass core and cladding
Operating wavelength	1300 nm
Core diameter[a]	62.5 ± 3.0 μm
Cladding diameter[b]	125 ± 3.0 μm
Numberical aperture[c]	0.275 ± 0.015
Minimum modal bandwidth[d]	500 MHz-km
Attenuation	1.75 dB/km at 1300 nm (maximum)

Note: From Ref. [11].
[a]Measured in accordance with EIA 455 FOTP 58, 164, 167, or equivalent.
[b]Measured in accordance with EIA 455 FOTP 27, 45, 48, or equivalent.
[c]Measured in accordance with EIA 455 FOTP 47 or equivalent.
[d]Measured in accordance with EIA 455 FOTP 51 or equivalent.

orange. All the parameters are specified and measured in accordance with the applicable industry standards as indicated.

21.3.7.2. Multimode Trunk Fiber Specification

Two multimode fiber types are supported for the trunk**.** The required optical parameters of both trunk fibers are specified in Table 21.6. Both fiber types conform to applicable European and U.S. industry standards [14–16]. All fiber parameters are specified and measured in accordance with the applicable industry standards as indicated.

21.3.8. ESCON Connector (Multimode)

The ESCON connector (illustrated in Fig. 21.5) is a ruggedized, two-ferrule connector that is polarized to prevent misplugging. The polarization is accomplished by beveling two corners of the connector as shown in Fig. 21.5. The

Table 21.6
Multimode Trunk Fiber Specifications.

Parameter	Specification
62.5/125 μm multimode fiber	
Fiber type	Graded index with glass core and cladding
Operating wavelength	1300 nm
Core diameter[a]	62.5 ± 3.0 μm
Core noncircularity	6% maximum
Cladding diameter[b]	125 ± 3.0 μm
Cladding noncircularity	2% maximum
Core and cladding offset	3.0 μm maximum
Numberical aperture[c]	0.275 ± 0.015
Minimum modal bandwidth[d]	500 MHz-km at ≤2 km
	800 MHz-km at >2 km and ≤3 km
Attenuation[e]	1.0 dB/km at 1300 nm
50/125 μm multimode fiber	
Fiber type	Graded index with glass core and cladding
Operating wavelength	1300 nm
Core diameter[a]	50 ± 3.0 μm
Core noncircularity	6% maximum
Cladding diameter[b]	125 ± 3.0 μm
Cladding noncircularity	2% maximum
Core and cladding offset	3.0 μm maximum
Numberical aperture[c]	0.200 ± 0.015
Minimum modal bandwidth	800 MHz-km
Attenuation[e]	0.9 dB/km at 1300 nm

Note: From Ref. [11].
[a]Measured in accordance with EIA 455 FOTP 58, 164, 167, or equivalent.
[b]Measured in accordance with EIA 455 FOTP 27, 45, 48, or equivalent.
[c]Measured in accordance with EIA 455 FOTP 47 or equivalent.
[d]Measured in accordance with EIA 455 FOTP 51 or equivalent.
[e]The attenuation is a typical value rather than a specification. Table 21.1 is the specification.

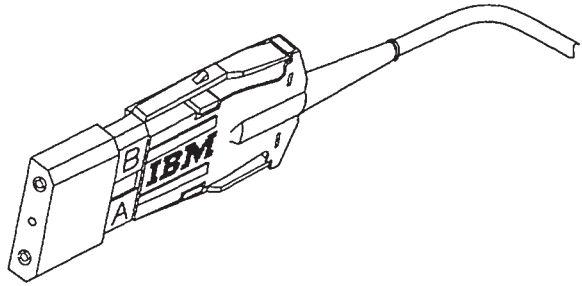


Figure 21.5 ESCON multimode duplex connector [12].

connector is a push-on type that is retained in the receptacle by two latches. The ferrules are protected from handling and dirt by a protective spring-loaded cap that retracts upon insertion in a receptacle, allowing the ferrules to be inserted into alignment sleeves. The ferrules are made of high-precision, stabilized zirconia ceramic. The ferrule ensures accurate alignment of the fiber to an LED, photodiode (PD), or fiber receptacle. The ends of the ceramic are polished into a convex shape to ensure that glass physical contact is always achieved in fiber-to-fiber connections. This feature ensures low-loss connections (typically less than 0.4 dB) and high return loss from the connection interface.

The color of the multimode connector is black. The connector also has bend radius-limiting boot at the connection of the fiber cable. This ensures that excessive bend stresses cannot be applied to the fiber at the connector.

21.4. SINGLE-MODE PHYSICAL LAYER

The single-mode physical layer follows the same approach as the multimode physical layer and defines a common set of specifications at the device interface that ensure interoperability [11]. The ESCON XDF link is composed of two simplex, point-to-point counterdirectional links encased in a duplex configuration. Both the optical and mechanical properties of the interface are specified as well as the cable plant.

The physical layer defines a set of specifications that ensure link performance up to a 20 km distance, without retransmission, using dispersion unshifted single-mode fiber. The data rate specification of 200 60.04 Mb/s is the same as multimode. A "1" bit corresponds to a light on condition.

All specifications are for worst-case operating conditions, including end-of-life conditions.

21.4.1. Single-Mode Output Optical Interface

The parameter specifications in Table 21.7 define the optical output interface for the light coupled into the single-mode fiber. The parameter specifications ensure the bit error rate does not exceed 10^{-15}. The light source is a 1300-nm semiconductor laser.

21.4.2. Single-Mode Input Optical Interface

The single-mode optical input interface requirements are given in Table 21.8. A loss-of-light function and operation is specified for link failure indication and diagnostic use. The design of the machine receiving this information determines how this state change information is utilized.

21.4.3. Single-Mode Fiber-Optic Cable Specification

The EXON XDF link is similar to the multimode link. It is composed of two unidirectional links assembled in a common cable and connector housing. The

Table 21.7
Single-Mode Optical Output Interface Specification.

Parameter	Minimum	Maximum	Unit
Average power into SMF[a]	–8.0	–3.0	dB
Center wavelength[a]	1261	1360	nm
Spectral width (rms)		7.7	nm
Rise time (t_r) (20–80%)[a,b]		1.5	ns
Fall time (t_f) (80–20%)[a,b]		1.5	ns
Eye window[a]	3.5		ns
Optical output jitter[c]		0.8	ns
Extinction ratio[a,d]	8		dB
Relative intensity noise (RIN_{12})[e]		–112.0	dB/km
t_r, t_f at optical path output[c,f]		2.8	ns

Note: From Ref. [11].
[a]Based on any valid 8B/10B code pattern, the measurement is made using 4-m single-mode jumper cable and only includes the power in the lowest order fundamental mode.
[b]The minimum frequency response bandwidth range of the optical waveform waveform detector shall be 100 kHz to 1 GHz.
[c]The optical output jitter includes both deterministic and random jitter. It is defined as the peak-to-peak time-histogram oscilloscope value (minimum of 3000 samples) using a 2^7-1 pseudo-random pattern or worst-case 8B/10B code pattern. The transmitter output light is coupled to a PIN photodiode O/E converter (e.g., Tektronix P6703A or equivalent) via a 3-m cable and jitter measured with a digital sampling oscilloscope [13].
[d]Measurement shall be made with a DC-coupled optical waveform detector that has a minimum bandwidth of 600 MHz and whose gain flatness and linearity are maintained over the range of the high and low optical power levels.
[e]The relative intensity noise is measured with a 12-dB optical return loss into the output interface.
[f]The maximum degradation in input interface sensitivity (from, e.g., jitter, mode hopping, and intersymbol interfacference) that can occur by using a worst-case link as specified in Table 21.2. The spectral width of the transmitter shall be controlled to meet this specification.

cable is terminated in a polarized duplex connector. The trunk cable is usually installed by professionals in high-count cables and physically distributed in distribution panels (see Section 21.5).

21.4.3.1. Single-Mode Jumper Cable Assembly

The SMF jumper cable assembly is a second-generation product that conforms to the ANSI Fibre Channel Standard [17]. The cable assembly is nonmetallic and uses aramid fiber as a strength member. The cable construction is illustrated in Fig. 21.6.

The single-mode jumper is only offered in 9/125 μm nondispersion-shifted fiber with the optical and mechanical specifications given in Table 21.9. The color

Table 21.8
Single-Mode Optical Input Interface Specifications.

Parameter	Minimum	Maximum	Unit
Sensitivity[a,b]		–28.0	dBm
Saturation level[a]	–3.0		dBm
Return loss[c]	12.5		dB
Acquisition time[c]		100	ns
LOL threshold[d]	–45	–32	dB
LOL hysteresis[e]	1.5		dB
Reaction time for LOL state change	0.25	5000	μs

Note: From Ref. [11].
[a]Based on any valid 8B/10B code pattern measured at, or extrapolated, 10^{-15} BER measured at center of eye. This specification shall be met with worst-case conditions as specified in Table 21.7 for the output interface and Table 21.2 for the fiber-optic link. This value allows for a 0.5-dB retiming penalty.
[b]A minimum receiver output eye opening of 1.4 ns at 10^{-12} should be achieved with a penalty not exceeding 1 dB.
[c]The measurement is made using 4-m single-mode jumper cable and only includes the power in the lowest order fundamental mode.
[d]The acquisition time is the time to reach synchronization after the removal of the condition that caused the loss of synchronization. The pattern sent for synchronization is either the idle character of an alternating sequence of idle and data characters.
[e]Required to avoid random transitions between LOL being active and inactive when input power is near threshold level.

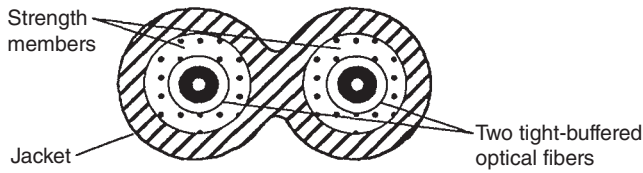


Figure 21.6 Single-mode jumper cable construction [12].

of the cable jacket is yellow. All parameters are specified and measured in accordance with the applicable industry standards as indicated.

21.4.4. ESCON XDF Duplex Connector

The single-mode ESCON XDF connector is the same as the FCS SC push-on, polarized (via keying) duplex fiber-optic connector (Fig. 21.7) [17].

The ferrules are also made of zirconia ceramic. The ends of the ferrules are polished to ensure physical contact when mated to another connector. For the

Table 21.9
Single-Mode Jumper Cable Specifications.

Parameter	Specification
Fiber type	Dispersion unshifted
Operating wavelength	1261–1360 nm
Mode field diameter[a]	9.0 ± 1.0 μm
Zero-dispersion wavelength[b]	1310 ± 10 μm
Dispersion[b]	6.0 ps/nm-km maximum
Cutoff wavelength(λ_c)[c]	1260 nm maximum
Attenuation[d]	0.8 dB/km maximum

Note: From Ref. [11].
[a]Measured in accordance with EIA 455 FOTP 164, 167, or equivalent.
[b]Measured in accordance with EIA 455 FOTP 168 or equivalent.
[c]Measured in accordance with EIA 455 FOTP 80 or equivalent.
[d]Measured in accordance with EIA 455 FOTP 78 or equivalent.

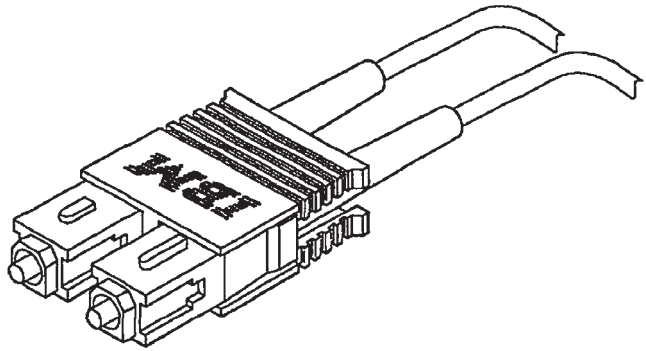


Figure 21.7 ESCON single-mode duplex connector [12].

single-mode version, one key is narrower than the other. It is mechanically retained in the duplex receptacle by latch arms that engage the connector upon plugging. Each fiber has its own subassembly that is compatible with the industry standard SC connector. The mating, external dimension, and interface requirements of the connector conform to the ANSI Fibre Channel Standard [17]. In order to improve usability, the connector housing is gray in color. This helps to differentiate the single-mode link from the multimode link. The connector also uses bend radius-limiting boots to minimize bending stress at the cable/connector interface.

21.4.4.1. Single-Mode Duplex Receptacle Specifications

The duplex receptacle for the XDF single-mode connector is specified in the Fibre Channel Standard FC-PH document. Any single-mode connector that is compliant with FCS will interoperate with the XDF single-mode receptacle.

21.5. PLANNING AND INSTALLATION OF AN ESCON LINK

The distance between ESCON-compatible equipment may vary from 4 m to 20 km and beyond when equipment (e.g., director) is used to retransmit the optical signal.

Fiber-optic cable planning is very similar to planning for a copper cable plant. The cables should be installed in troughs or conduits, which may be run in the ceiling or under the floor. For a small installation, where the quantity of equipment is small and the equipment pieces are located close to each other, the cable plant will likely consist of jumper cables only. Cable specifications are given in Tables 21.10 and 21.11.

Table 21.10

Multimode Jumper Physical and Environmental Specifications.

Parameter	Specification
Physical	
Jacket material	PVC
Jacket outside diameter	4.8 mm
Weight (for information only)[a]	20 g per meter
Installation tensile strength	1000 N maximum
Minimum bend radius (during installation)	4.0 mm; 5 s maximum at 400 N
Minimum installed bend radius	
No load	12 mm
Long-term residual	25 mm at 89 N
Flammability	Underwriters laboratory UL-1666 ONFR (Optical Fiber Nonconductive; Plenum UL-910 is also acceptable)
Crush resistance	500 N/cm maximum
Maximum unsupported vertical rise	100 m
Environmental	
Operating environment	Inside building only
Operating temperature	10–60°C
Operating relative humidity	5–95%
Storage and shipping	–40 to 60°C
Lightning protection	None required
Grounding	None required

Note: From Ref. [21].

[a]Cable only; connectors not included.

Table 21.11
Single–Mode Jumper Physical and Environmental Specifications.

Parameter	Specification
Physical	
Jacket material	PVC
Jacket outside diameter	Zip cord
Weight (for information only)[a]	20 g per meter
Installation tensile strength	1000 N maximum
Minimum bend radius (during installation)	50 mm; 5 s maximum at 100 N
Minimum installed bend radius	
No load	30 mm
Long-term residual	50 mm at 80 N
Flammability	Underwriters laboratory UL-1666 ONFR (Optical Fiber Nonconductive; Plenum UL-910 is also acceptable)
Crush resistance	899 N/cm maximum
Maximum unsupported vertical rise	100 m
Environmental	
Operating environment	Inside building only
Operating temperature	0–60°C
Operating relative humidity	8–95%
Storage and shipping	–40 to 60°C
Lightning protection	None required
Grounding	None required

Note: From Ref. [21].
[a]Cable only; connectors not included.

Light propagation occurs in a single direction within a fiber. In an ESCON jumper cable (for either multimode or single mode), light enters the "B"-labeled port and exits the "A"-labeled port of the cable assembly (Fig. 21.8). The transceiver output is always coupled to the B port, and the transceiver input is always coupled to the A port. The fiber crossover is designed into the jumper cable assembly, and, by virtue of the polarized connector, the transceiver output is always connected to the corresponding transceiver input at the other end of the cable.

The design of the trunk cable plant must be done to account for the fiber crossover. Only an odd number of crossovers will result in a viable link. In general, a link contains two jumper cables, which requires the trunk fiber to be crossed in order to ensure link continuity. This configuration is illustrated in Fig. 21.9.

The physical location and quantity of equipment are the primary factors affecting the cable layout. The application environment may vary from a single-room configuration with limited equipment to a multicampus environment with multiple buildings and floors within a building.

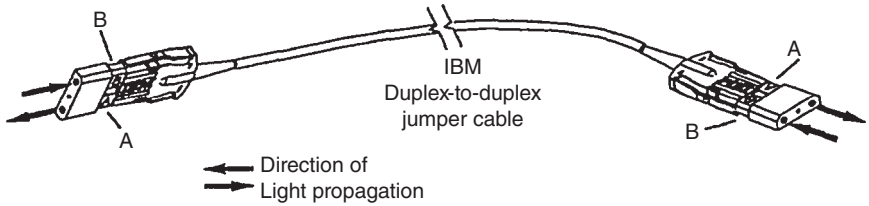


Figure 21.8 Direction of light propagation in an IBM jumper cable (multimode depicted) [20].

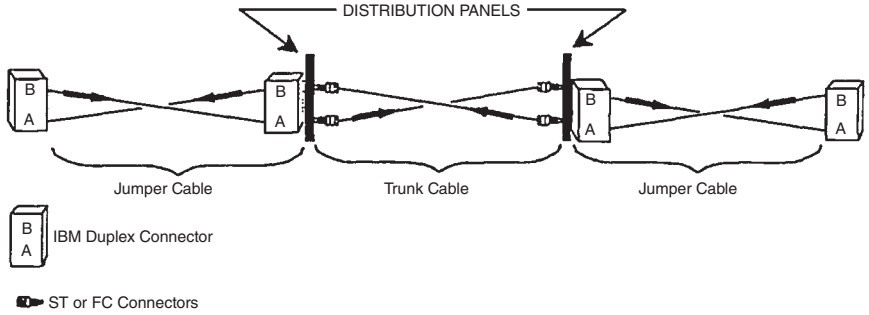


Figure 21.9 Direction of light propagation in a link with a trunk [20].

The example illustrated in Fig. 21.10 is a single campus environment with data processing equipment in two buildings. The point of egress of each building is at a "building interface panel." This panel provides a common access point for each building. The trunk cable (multimode or single mode) between the buildings will be determined by the actual "run" distance of the intended cable. In general, single mode is chosen for distances greater than 3 km. The trunk cable should be a rugged outdoor cable type that can be installed in an underground conduit (preferred) or strung on utility poles. Extra count trunks should be considered for expansion.

Within a building, distribution panels are optimally located. Both building trunks and jumpers are used to interconnect distribution panels. In general, jumpers connecting to distribution panels should be kept as short as possible. Distribution panels should be optimally located with multifiber trunks interconnecting the distribution panels.

The flexibility and congestion relief inherent with distribution panels requires additional termination hardware and increased link loss due to the inclusion of additional connections and splices. This must be considered in the loss budget analysis (Section 21.6) to ensure the loss budget requirements are not violated.

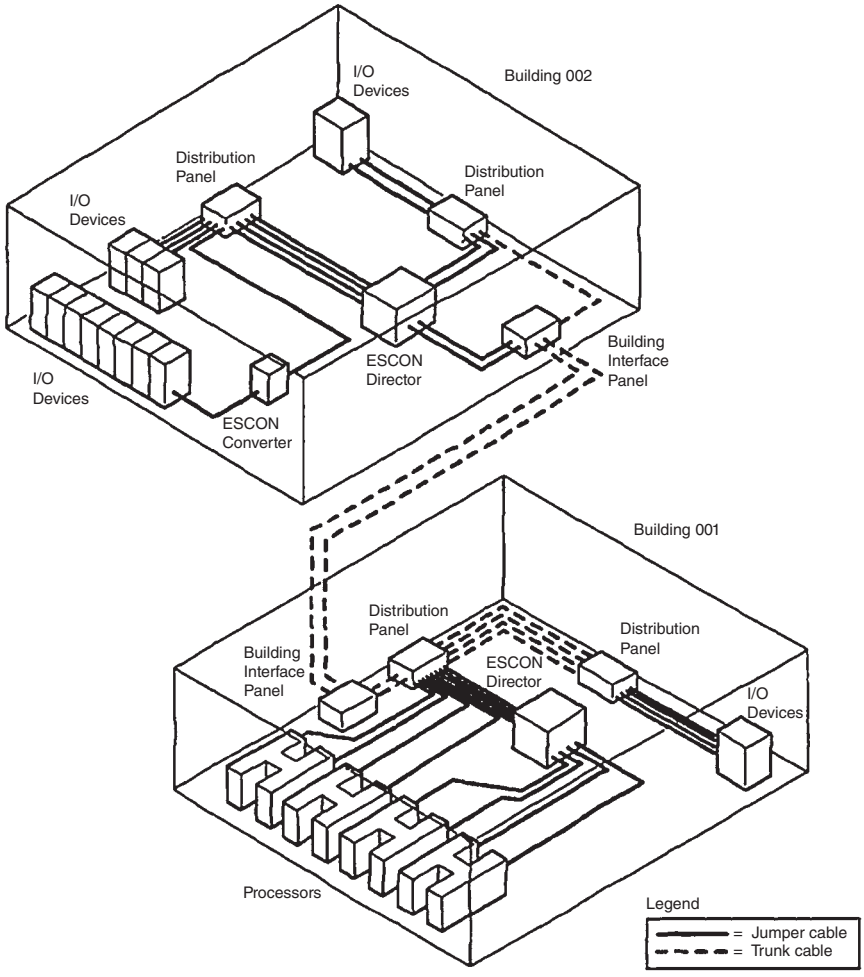


Figure 21.10 Example of a single campus link environment between two buildings [20].

If the two buildings were located on different campuses, then right-of-way access is required from the owner of the land. Leasing of "dark fiber" may also be a possibility. The negotiation, however, for right-of-way, leasing, and so on, may greatly extend the total planning and installation time. This must be considered in the overall project schedule.

For distance requirements beyond 20 km (XDF feature maximum), channel extenders may be used such as the IBM 9036 ESCON Remote Channel Extender. The IBM 9036 is a repeater that retransmits the optical signal to achieve an

additional 20-km distance. Various types of fiber-optic multiplexers may also be used to increase the distance up to 100 km or more (see Chapter 15).

The capability of the base ESCON channel architecture and implementation (point-to-point multimode fiber link is limited to 3 km) is designed to support a total channel distance of 9 km or greater between the host and the device, by using intervening directors and/or channel extenders. The distance capability of ESCON is an important attribute of the technology and has been utilized extensively. This is a significant improvement over the prior channel distance of 400 feet, using parallel copper technology. Also, the use of the XDF link, available between directors, and single-mode fiber extenders can extend the total channel distance to 60 km. This capability provides improved configuration flexibility to optimize the placement of ESCON-capable data processing equipment with regard to installation and facilities constraints as well as enhanced disaster recovery. Data processing equipment can now be placed in nearby buildings, as required by the customer.

ESCON performance, in general, is relatively constant up to 9 km (degrading by 10% at 9 km) and begins to droop substantially thereafter. This distance capability allows high-bandwidth devices (e.g., DASD) to be placed anywhere within a "typical" campus environment while maintaining the high-performance capability of the ESCON channel. In addition, some low-speed devices (e.g., printers) may be placed as far as 60 km from the host and will perform acceptably.

The performance droop results from effects which include a large number of acknowledgments for each data transfer. Some recent applications, such as the IBM Virtual Tape Server (VTS), can run over an ESCON physical layer without requiring these handshake acknowledgments; these applications can maintain their data throughput up to significantly longer distances (100 km or more).

For large installations within a single building or data center, the use of direct attach trunk cables has become popular [20]. Multifiber ribbon (12 fibers per ribbon) trunk cables are directly attached to an IBM eServer zSeries as illustrated in Fig. 21.12. For large ESCON channel count servers, direct-attached trunks offer the least congested cabling approach.

The trunk may contain as many as 12 ribbons (144 fibers total). Each ribbon is terminated in a 12-fiber multifiber terminated push-on (MTP) connector. The small size of the connectors easily enables six MTPs to be terminated inside the server frame in a coupler bracket. Several brackets can be installed in a machine. A harness is used inside the machine to convert the 12 fibers (representing six ESCON channels) of a ribbon to six ESCON connectors for insertion into the appropriate transceiver.

Once the harnesses are installed in the machine, all fiber connects and disconnects can be done using the trunk cable MTP connectors. The harnesses remain inside the machine, plugged into the individual ESCON or coupling facility ports, whereas the trunk cables can be quickly removed. The machine

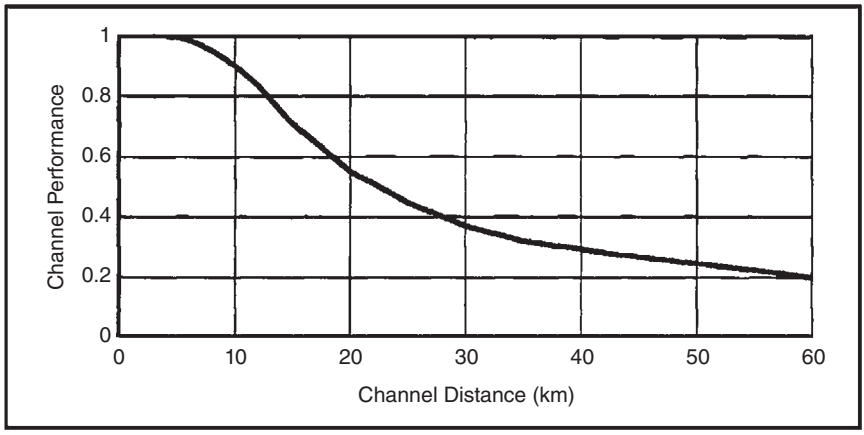


Figure 21.11 Representative ESCON channel performance as a function of distance (courtesy of IBM Corporation).

can then be relocated within the data complex because the trunk cables are easily rerouted to the new location and replugged into the harnesses. This greatly reduces the time spent unplugging, rerouting, and plugging the individual fiber-optic jumper cables back into the machine ports. MTP-to-MTP multiribbon trunks can be used throughout the complex to directly connect equipment together as well as equipment to distribution panels. For these installations the use of jumper cables is minimized. This cabling approach is used for both multimode and single-mode versions. IBM Availability Services offers this approach as Fiber Transport System I11 (FTS-111) and is also available to do the planning, design, and installation of this interconnection system.

21.6. LOSS BUDGET ANALYSIS

The cable plant loss (C,) of an ESCON link is a critical parameter and is determined by the link length, fiber attenuation, number of connections and jumper cable loss, jumper/trunk core diameter mismatch splices, and the like [22]. These parameters are greatly influenced by the cable plant layout and quality of components used. As discussed in Section 21.3.3, the required optical cable plant loss is established by the available power as determined by Eq. (21.1). The cable plant optical link loss must be less than or equal to the available power to ensure the link design performance. This is true for both multimode and single-mode links.

21.6.1. Multimode Cable Plant Link Loss

The multimode link loss is determined by statistically adding all the individual loss elements in the cable plant. A good approximation is that the individual loss distributions are Gaussian and can be described with a mean value and variance.

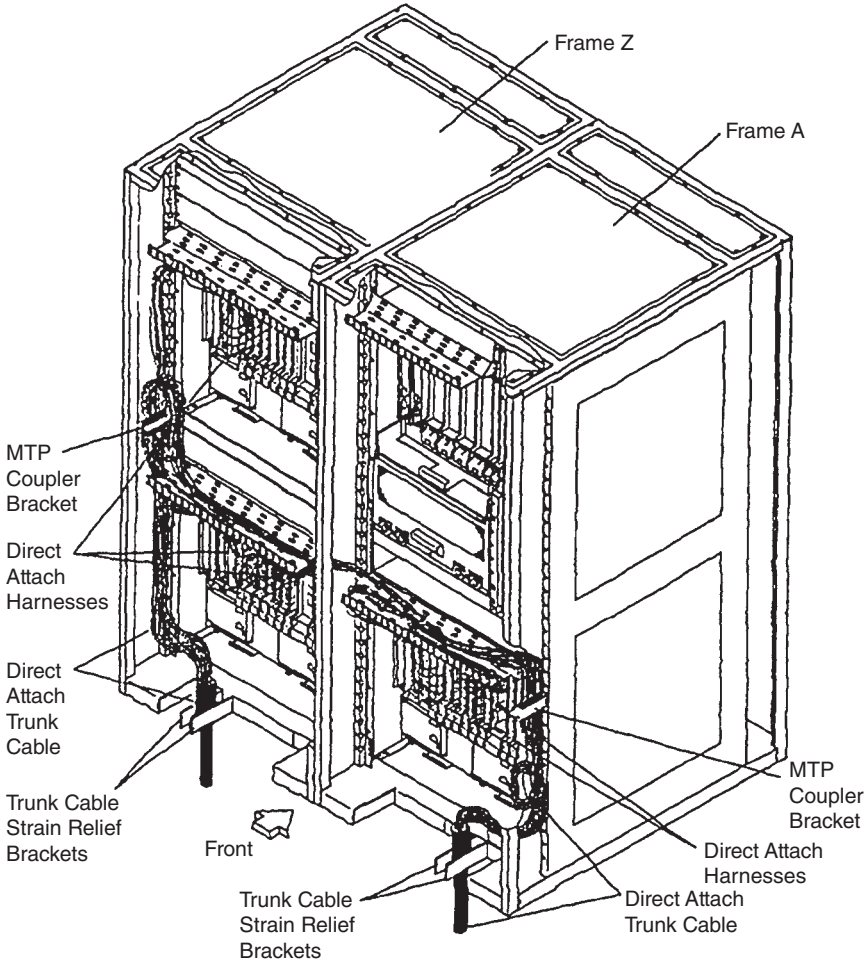


Figure 21.12 IBM's direct-attach trunk system in an eServer zSeries [20].

The means and variances of all the individual elements are summed together to yield an overall cable plant mean (M_c) and variance (σ_c^2), where σ_c is the cable plant loss standard deviation. The total cable plant loss is given by

$$C_t = M_c + 3\sigma_c \tag{21.2}$$

Representative loss values for the individual elements can be obtained from component manufacturers or from a professional fiber-optic network planner. Table 21.12 contains typical loss distributions for the elements in a cable plant. Table 21.12 can be used as a guide if more accurate data are not available; however, it is preferred to use data from the cable plant installer or component

Table 21.12
Typical Optical Component Loss Values.

Component	Description	Fiber Core Size (μm)	Mean Loss (dB)	Variance (dB^2)
Connector[a]	Physical contact	62.5–62.5	0.40	0.02
		50.0–50.0	0.40	0.02
		9.0–9.0[b]	0.35	0.02
		62.5–50.0	2.10	0.12
		50.0–62.5	0.00	0.01
Splice	Mechanical	62.5–62.5	0.15	0.01
		50.0–50.0	0.15	0.01
		9.0–9.0[b]	0.15	0.01
Splice	Fusion	62.5–62.5	0.40	0.01
		50.0–50.0	0.40	0.01
		9.0–9.0[b]	0.40	0.01
Cable	IBM MMF jumper	62.5	1.75[c]	NA
	IBM SMF jumper	9.0	0.8[c]	NA
	Trunk	62.5	1.00[c]	NA
	Trunk	50.0	0.90[c]	NA
	Trunk	9.0	0.50[c]	NA

Note: From Ref. [21].
[a]The connector loss value is typical when attaching identical connector types. The loss can vary significantly if attaching different connector types.
[b]Single-mode connectors and splices must meet a minimum return loss of 28 dB.
[c]Actual loss value in dB/km.

manufacturer. In general, component losses (e.g., fiber attenuation, and connection loss) are measured using an equilibrium mode distribution in the fiber. In order to account for the optical loss of a link associated with mode redistribution in multimode fibers, the following loss should be included in a link:

- 1.5 dB for 50/125 μm trunk fiber, or
- 1.0 dB for 62.5/125 μm trunk fiber

As a multimode link, a typical configuration might be two jumper cables with a combined length of 90 m, 1.5 km of 50/125 μm trunk (800 MHz/km bandwidth), two physical contact connectors (one 62.5–50 μm and the other 50–62.5 μm), and six splices. The element loss values used are from Table 21.12. In this case the total cable plant link loss is 7.3 dB.

21.6.2. Single-Mode Cable Plant Link Loss

This section describes the calculation of the loss in a single-mode link. The approach and procedure are the same as those in the multimode link. The

individual element losses are statistically added to yield a total link loss mean and variance value. Equation (21.2) is used to calculate the cable plant loss. In the case of a single-mode fiber, no equilibrium mode distribution loss occurs, but there is a small loss associated with mode field diameter mismatches (excess connector loss). This loss adds to a 0.5-dB total (two connections) and is small enough to be included in the loss budget for all configurations (whether present or not). A typical single-mode link could consist of two jumper cables with a combined length of 210 m, 19.76 km of 9/125 μm trunk, two physical contact connections, and two mechanical splices.

21.6.3. SBCON ANSI Standard ESCON

ESCON has proven to be an excellent communication link for the data processing environments. IBM, with the support and encouragement of other companies in the information industry, has requested that ANSI sanction an official work effort to make ESCON an industry standard. The activity was formally approved by ANSI in 1995, and a draft standard SBCON has been established [25]. The SBCON standard was approved in 1997. The information contained in this chapter is consistent with this standard.

21.6.4. ESCON Technology Update

More recent ESCON I/O card designs feature a high-port density card as illustrated in Fig. 21.13. The ESCON card density has been improved by 4x with the use of state-of-the-art optical and silicon technologies. The ESCON transceiver has been changed to an industry standard, small form factor (SFF) transceiver. This upgrade enables the placement of 16 transceivers along the card edge (shown on the right side in Fig. 21.13). The four ASIC modules directly behind the transceivers contain the ESCON channel logic, serializer, and deserializer functions. Each module contains four ESCON independent channels. The remaining components contain support functions. The ESCON optical specifications and link performance have not been changed. The link is functionally fully compatible with existing ESCON links.

The use of the SFF transceiver necessitated a change of the ESCON connector from the original design (Fig. 21.5) to a smaller MT-RJ fiber-optic connector. The MT-RJ uses a single plastic molded ferrule instead of two ceramic ferrules. It is an axial push/pull connector with a single latch. The connector conforms to the ANSI/TIA/EIA 604-12 (FOCUS 12) fiber-optic connector standard [26]. An MT-RJ fiber-optic jumper cable is illustrated in Fig. 21.14.

The smaller connector and cable size improves the handling and flexibility of the jumper cable which, in turn, improves jumper cable installation. The new jumper cable is optically and functionally compatible with existing links and cables. However, a connector conversion is necessary to attach a new transceiver

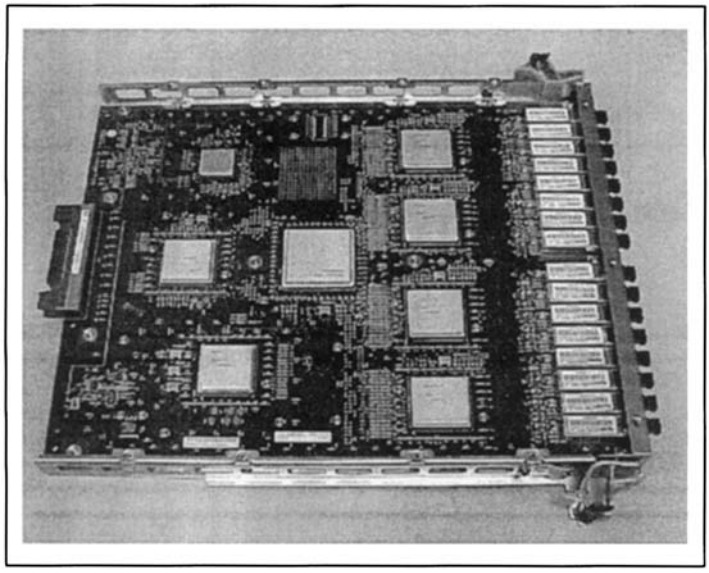


Figure 21.13 High-density ESCON I/O card (courtesy of IBM Corp.)

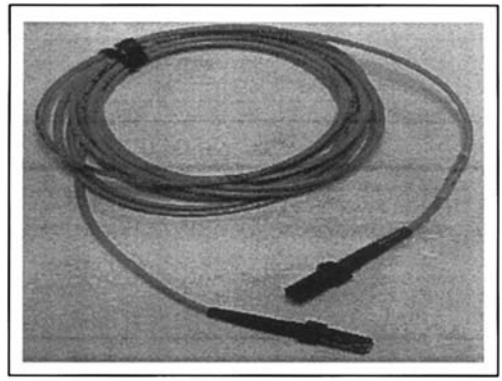


Figure 21.14 MT-RJ ESCON multimode jumper cable (courtesy of IBM Corp.).

or cable to an existing cable infrastructure. A cable conversion kit from original ESCON to MT-RJ is available from IBM to enable this conversion without changing the existing cable infrastructure. New installations should design a cable infrastructure compatible with the ESCON MT-RT cable system.

There has always been a limited interest in non-standard implementations of ESCON and other protocols. One example is the parallel ESCON transceiver,

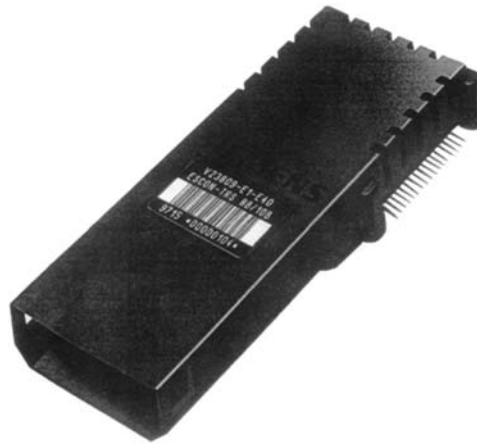


Figure 21.15 ESCON parallel transceiver.

developed by Siemens for customers interested in using many ESCON channels in a small space. This module, shown in Figure 21.15, also contains circuitry functions such as serializing/deserializing of 10 single-ended parallel data inputs and outputs, each of 20 Mbit/s, as well as clock transfer and synchronization. This solution reduces the board space required for ESCON channels by roughly a factor of 4 compared to discrete solutions (e.g., with serial ESCON/SBCON transceivers).

ACKNOWLEDGMENTS

I wish to thank the International Business Machines Corporation for allowing the extensive use of copyrighted material in this chapter. I also wish to thank the many individuals (too numerous to list here) both within and outside of IBM whose hard work and dedication helped make ESCON a reality. In particular, I wish to thank M. Lewis and J. Quick for the hardware photographs.

REFERENCES

1. IBM Corp. *Introducing enterprise systems connection* (IBM Document No. GA23–383). Mechanicsburg, Pa.: IBM Corp.
2. Calta, S. A., J. A. deVeer, E. Loizides, and R. N. Strangwayes. 1992. Enterprise systems connection architecture—System overview. *IBM J. Res. Dev.* 36:535–551.
3. Georgiou, C. J., T. A. Larsen, P. W. Oakhill, and B. Salimi. 1992. The IBM enterprise systems connection (ESCON) director: A dynamic switch for 200 Mb/s fiber optic links. *IBM J. Res. Dev*. 36:593–616.
4. Elliott, J. C., and M. W. Sachs. 1992. The IBM enterprise systems connection architecture. *IBM J. Res. Dev*. 36:577–592.

5. Flanagan, J. R., T. A. Gregg, and D. E. Casper. 1992. The IBM enterprise systems connection (ESCON) channel—A versatile building block. 1992. *IBM J. Res. Dev.* 36:617–632.
6. IBM Corp. Enterprise systems architecture/390 EXON I/O interface (IBM Document No. SA22–7202, Chap. 2). Mechanicsburg, Pa.: IBM Corp.
7. IBM Corp. Enterprise systems architecture/390 ESCON 1/0 interface (IBM Document No. SA22–7202, Chap. 8). Mechanicsburg, Pa.: IBM Corp.
8. IBM Corp. Enterprise systems architecture/390 ESCON I/O interface (IBM Document No. SA22–7202, Appendix). Mechanicsburg, Pa.: IBM Corp.
9. Aulet, N. R., D. W. Boerstler, G. DeMario, E. D. Ferraiolo, C. E. Hayward, C. D. Heath, A. L. Huffman, W. R. Kelley, G. W. Peterson, and D. J. Stigliani, Jr. 1992. IBM Enterprise systems multimode fiber optic technology. *IBM J. Res. Dev.* 36:553–577.
10. DeCusatis, C. 1995. Data processing systems and optoelectronics. In *Optoelectronics for Data Communication*, R. C. Lasky, U. L. Osterberg, and D. J. Stigliani (eds.), Chapter 6. New York: Academic Press.
11. IBM Corp. Enterprise Systems Architecture/390 ESCON I/O interface physicallayer (IBM Document No. SA23-0394). Mechanicsburg, Pa.: IBM Corp., courtesy of International Business Machines Corporation.
12. IBM Corp. Planning for fiber optic channel links (IBM Document No. GA23-0367, Chap. 1). Mechanicsburg, Pa.: IBM Corp., courtesy of International Business Machines Corporation.
13. Gregurick, V. 1992. Fiber optic transmitter measurement procedure, IBM Engineering Specification 49G3489. East Fishkill, N.Y.: IBM Corp.
14. Electronics Industry Association/Telecommunications Industry Association Commercial Building Telecommunications Cabling Standard (EIA/TIA-568-A).
15. Electronics Industry Association. Detail specification for 62.5 μm core diameter/l25 μm cladding diameter class la multimode. Graded index optical waveguide fibers (EIA/TIA-492AAAA). New York: Electronics Industry Association.
16. International Telecommunications Union. Characteristics of a 50/125 μm multimode graded index optical fibre cable (CCITT Recommendation G.651).
17. Fiber Channel Standard. Physical layer and signaling specification. American National Standards Institute X3T9/91-062, X3T9.3/92-001 (rev. 4.2).
18. Electronics Industry Association Detail specification for class IVA dispersion unshifted single-mode optical wavelength fibers used in communications systems (EIA/TIA-492BAAA). New York: Electronics Industry Association.
19. International Telecommunications Union. Characteristics of single-mode optical fibre cable (CCI'IT Recommendation G.652).
20. IBM Corp. Planning for fiber optic channel links (IBM Document No. GA23-0367, Appendix A). Mechanicsburg, Pa.: IBM Corp., courtesy of International Business Machines Corporation.
21. IBM Corp. Planning for fiber optic channel links (IBM Document No. GA23-0367, Chap. 2). Mechanicsburg, Pa.: IBM Corp., courtesy of International Business Machines Corporation.
22. IBM Corp. Planning for fiber optic channel links (IBM Document No. GA23-0367, Chaps. 3 and 4). Mechanicsburg, Pa.: IBM Corp., courtesy of International Business Machines Corporation.
23. Light launch conditions for long-length graded-index optical fiber spectral attenuation measurements (EIA/TIA-455-50A), available from Electronics Industry Association, New York.
24. IBM Corp. Maintenance information for enterprise systems connection links (IBM Document No. SY27-2597). Mechanicsburg, Pa.: IBM Corp.
25. American National Standards Institute. Single-byte command code sets CONnection architecture (SBCON). Draft ANSI StandardX3Tl 1/95-469 (rev. 2.2).
26. American National Standards Institute. ANSI/TIA/EIA 604-12 (FOCUS 12) Fiber Optic Connector Intermateability Standard, Type MT-RJ.

Case Study Calculating an ESCON Optical Link Budget

Courtesy of IBM Corporation

Application: Design an optical link budget to support FICON transmission on a 10-km single-mode link with a target BER of 10e-12.

Description: The following example discusses the design of a data communication link using single-mode optical fiber operating at 1300 nm. The application is remote backup of data from a hard disk to a tape drive located in a building 10 km away. Because of systemwide fault tolerance, the link has a target BER of 10e-12. The link adheres to an industry standard data rate of 200 Mbit/s. The transmitter is guaranteed to have a worst-case output of –9.0 dBm, accounting for temperature variation and end of life; the worst-case receiver sensitivity is –26.0 dBm. Optical fiber is chosen with parameters shown in the attached tables, which also give a complete specification for the transmitter and receiver. Note that the link center wavelength is sufficiently close to the zero-dispersion wavelength of the fiber that the dispersion penalty is kept to 0.2 dB worst case. The transmitter is class 1, so we can neglect nonlinear effects in the link; connector return loss is held low enough so that multipath effects are not a concern either. Using the manufacturer's specification for RIN and assuming $g = 0.5$, we can estimate the RIN power penalty to be 0.7 dB, which should be tolerable. The wavelength-dependent attenuation of the fiber was selected so that the penalty would be tolerable at the required 10-km distance. For this distance, we have chosen a single-mode link, so modal noise is not a concern provided we use sufficiently low-loss splices and connectors to minimize modal noise. Although mode partition noise should be verified experimentally, we estimate that the power penalty is less than 1.5 dB

(note that k < 1, so the approximate model is valid). This gives us a total available power of

$$\textbf{Pavail} = \text{Tx(output)} - \text{Rx(sens)} - \text{Penalties} = 14\,\text{dB}$$

Transmitter	
Power output	–3 to –9 dBm
Extinction ratio	6 dB
RIN	–112 dB/Hz
Center wavelength	1270–1340 nm
Spectral width	6 nm
Receiver	
Sensitivity	–27 dB
Saturation	–3 dBm
Fiber	
Zero-dispersion wavelength	1310 nom. 1295–1322 (range) nm
Zero-dispersion slope	0.095 ps/nm-km
Cutoff wavelength	280 nm
Attenuation	0.5 at 1310 nm dB/km
Max. attenuation delta (1270–1340)	0.06 dB/km

Next, we compute the installation loss. The link contains one distribution panel near the source end and uses previously installed connectors. We will consider a link with four connectors (mean loss = 0.35 dB; variance = 0.06 dB), two fusion splices (mean loss = 0.15 dB; variance = 0.01 dB), and a 20-km link with 240 m of jumper cable (loss = 0.8 dB/km for the jumpers and 0.5 dB/km for the trunk). If we perform a statistical calculation on the installed budget for connectors, splices, and cable transmission loss, the mean installed loss is 11.75 dB with a variance of 0.26 dB. If we use the industry-standard "three sigma" design, then the calculated installation loss is 13.5 dB. This allows us to operate the link with 0.5 dB to spare, the difference between the available power and installation loss budgets. A conservative design should always allow some link safety margin, as we have done, to allow for the many assumptions used in deriving the various link performance models we have used. Of course, there are other possible ways to compute the link budget; if we had obtained the statistics for all the link parameters, we could perform a Monte Carlo simulation for the entire link that might yield a different link budget margin. It should be apparent that link design requires not only the latest valid perfomance models but also a good amount of engineering judgment in how best to apply these tools. Also, real-world problems are often not as clear-cut as this example; a link designer must be able to manage the tradeoffs involved in frequently working with incomplete information. Still, this example should illustrate some of the typical steps required to perform link planning and installation.

Sources	Typical Loss
Fiber attenuation	0.5 dB/km
Connector loss	0.25–0.50 dB
Splice loss	0.15–0.40 dB
Transmitter/receiver end-of-life degradation	0.5–1.0 dB

Penalty	Typical at 20 Km, 200 Mb/s(dB)
Chromatic dispersion	1.2
Mode partition	1.5
RIN	0.8
Multipath	0.25

This page intentionally left blank

22

Enhanced Ethernet for the Data Center[1]

Casimer DeCusatis
IBM Corporation

This chapter will provide an overview of Ethernet protocols, including both classic versions and more recent updates targeted for data center applications. Ethernet was originally developed in the early 1970s by Robert Metcalfe and David Boggs at the Xerox PARC research labs (and patented in 1978; see U.S. patent 4063220). This version was originally called Experimental Ethernet and was based, in part, on the wireless protocol called ALOHAnet. Since then, there have been many forms of this interface based on various IEEE standards (see Appendix E). In practice, the technical community has generally accepted the name "Ethernet" in reference to any of these standards. However, it should be noted that there can be significant technical differences between these interfaces, particularly at higher data rates (10–100 Gbit/s). Care must be taken not to assume that any protocol called Ethernet, standardized or otherwise, will necessarily refer to the same inexpensive technology that has grown to dominate local area networks, or that the components and specifications imply that all Ethernet-based standards will interoperate or have similar attributes. In fact, the Ethernet protocol has continually evolved and grown over time, to the point where it has recently been suggested as a candidate for the convergence of many other protocols in future systems.

Fundamentally, Ethernet includes a large family of frame-based networking protocols that operate over different types of media at different data rates. We will describe implementations of Ethernet using both conventional copper cables

[1]Portions of this chapter were adapted from R. Thapar, Chapter 15 in the second edition of the *Handbook of Fiber Optic Data Communications*. Review and comments from Scott Kipp and from Cisco Systems are gratefully acknowledged.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

and optical fiber, as well as the many modifications that have been proposed to enhance the robustness of this protocol for metropolitan area network (MAN) and data center applications.

22.1. INTRODUCTION TO CLASSIC ETHERNET

The early version of this protocol relied on a local area network (LAN) with multiple computers communicating over a shared coaxial cable; the common cable was likened to the old concept of the "ether" as a signal propagation media, from which Ethernet protocol takes its name. Much of the original Ethernet standard dealt with mechanisms to minimize collisions between data packets from different users by randomizing the times when a given computer could access the network. Over the years, Ethernet has been standardized by the Institute of Electrical and Electronic Engineers (IEEE) as IEEE 802.3 and has evolved into a much richer communications protocol, largely replacing competing LAN standards in conventional systems.

A mapping between the standard 7-layer Open Systems Interconnect (OSI) model and the Ethernet protocol stack is shown in Figure 22.1. Ethernet defines a number of possible physical layers and physical medium-dependent (PMD) sublayers, which will be discussed later in this chapter. An optional auto-negotiation layer provides interoperability with different data rate options.

At the Media Access Control (MAC) layer, a scheme known as carrier sense multiple access with collision detection (CSMA/CD) determines how multiple computers share a communications channel. This protocol required each computer to listen to the shared channel in order to determine its availability for

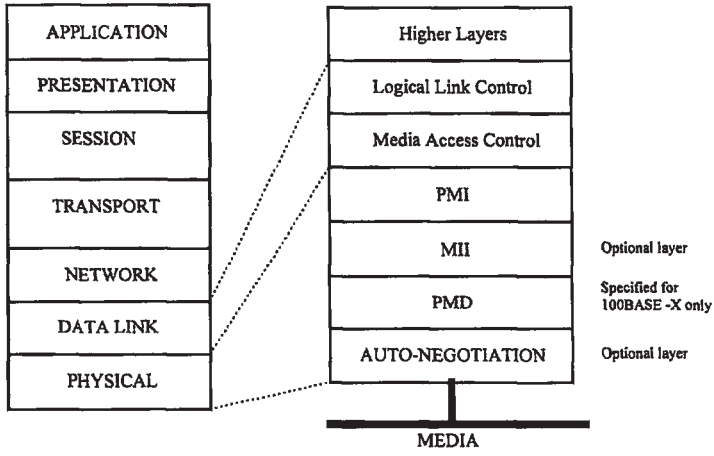


Figure 22.1 Mapping Ethernet stack to the 7-layer OSI model.

transmission. When the channel is silent, transmission can begin; if two computers attempt to transmit data at the same time, a collision occurs. Both computers then stop their transmission and wait for short, random periods of time (typically a few microseconds). In this way, it is unlikely that both computers will choose the same time to begin transmitting again, thus avoiding another collision. When there is more than one failed transmission attempt, exponentially increasing wait times are invoked as determined by a truncated binary exponential backoff algorithm. This approach is simpler than competing technologies such as token ring or token bus, which involve a set of rules for circulating a token on the network and only allow the current token owner access to the channel. However, this classic approach to shared media Ethernet also suffers from various weaknesses. For example, simply using a shared channel implies that the optical or copper cable is a potential single point of failure. In larger, extended distance networks, damage to a single portion of the cable plant could result in the entire network failing. Furthermore, since all communication takes place on a single channel, information intended for one computer can potentially be received by all other computers on the network; this poses a security risk. Finally, since network bandwidth is shared, performance can become quite slow under certain conditions.

Ethernet data packets can be organized into several types of standard frame formats. The most commonly used today is the Ethernet Version 2, or Ethernet II, frame, shown in Fig. 22.2. This is also sometimes known in the early literature as a DIX frame, named after the founding companies DEC, Intel, and Xerox, which helped pioneer this protocol. This frame is often used directly by Internet Protocol (IP), which consititutes the vast majority of applications for Ethernet in current systems.

The frame contains start bits, sometimes called the preamble, a logical link control (LLC), and a trailing parity check called the frame check sequence (FCS) or cyclic redundancy check (CRC). Note that these bits are removed by the Ethernet adapter before being passed on the the network protocol stack, so they may not be displayed by packet sniffers. The preamble consists of a 56-bit pattern of alternating 1 and 0 bits, which allow devices on the network to detect a new incoming frame. This is followed by a Start Frame Delimiter (SFD), an 8-bit value (10101011) marking the end of the preamble and signaling the start of the actual frame. The frame contains two MAC addresses for the packet destination and source. Following these is a subprotocol label field called the EtherType, which relates to the Maximum Transmission Unit (MTU), or the size of the largest

Preamble / SFD / destination MAC address / source MAC address / EtherType / Payload…./ FCS

Figure 22.2 Ethernet type II frame.

allowed packet, in bytes. Note that MTU is defined by the Ethernet standard, while other point-to-point serial links may negotiate the MTU as part of the connection initialization. There is a tradeoff in selecting the MTU; higher values allow greater bandwidth efficiency. However, larger packets can also cause congestion at slower speed interfaces, increasing the network latency (for example, a 1500-byte packet can occupy a 14.4-kbps modem for about 1 second).

Early Ethernet links were more prone to errors and tended to operate at lower data rates; thus the MTU was set to a maximum value of around 1500 bytes. If a packet was corrupted during transmission, recovery would only require the retransmission of the last 1500 bytes. However, as link error rates were reduced over time and data rates increased, it became possible to transfer and process larger data frames with lower server utilization. For these reasons, two basic EtherType fields have been standardized. Values of the EtherType field between 0 and 1500 indicate the use of the classic Ethernet format with an MTU up to 1518 bytes, while values of 1536 or greater indicate the use of a new larger frame format (the Q-tag for VLAN and priority data in IEEE 802.3ac extends the upper limit to 1522 bytes). While the larger MTU can refer to any frame greater then 1518 bytes in length, the most common implementation is a 9000-byte packet called Ethernet Jumbo Frames. There is a minimum packet size of 64 bytes; data units less than this size will be padded until the data field is 64 bytes large.

The terminology for a packet and a frame are often used interchangably in the literature, although this is not technically correct. The IEEE 802.3 and ISO/IEC 8802-3 ANSI standards define a MAC sublayer frame including fields for the destination address, source address, length/type, data payload, and FCS. The Preamble and Start Frame Delimiter (SFD, an 8-bit value marking the end of the preamble) are usually considered a header to the MAC frame, and the combination of a header plus a MAC frame constitutes a packet. IEEE 802.3 defines the 16-bit field after the MAC addresses as a length field with the MAC header followed by an IEEE 802.2 LLC header. It is thus possible to determine whether a frame is an Ethernet II frame or an IEEE 802.3 frame, allowing the coexistence of both standards on the same physical medium. All 802.3 frames have an IEEE 802.2 LLC header. By examining this header, it is possible to determine whether it is followed by a subnetwork access protocol (SNAP) header.

The LLC header includes two additional 8-bit address fields, called service access points, or SAPs. When both source and destination SAP are set to the value 0xAA, the SNAP service is requested. The SNAP header allows EtherType values to be used with all IEEE 802 protocols, as well as supporting private protocol ID spaces. In IEEE 802.3x-1997, the IEEE Ethernet standard was changed to explicitly allow the use of the 16-bit field after the MAC addresses to be used as a length field or a type field.

Other types of frames may be encountered in an Ethernet network. For example, Novell has developed a proprietary version of the IEEE 802.3 frame without

a standard LLC header. A frame may also contain optional fields which identify its quality of service level (IEEE 802.1p priority) or indicate which virtual local area network (VLAN) contains the link (IEEE 802.1q tag). While these frame types may have different formats and MTU values, they can coexist on the same physical link.

22.1.1. Auto-Negotiation

Auto-negotiation is the process that enables two devices, sharing a link segment, to communicate necessary information with one another in order to interoperate, taking maximum advantage of their abilities. The basis of auto-negotiation is a modified 10BASE-T link integrity pulse sequence as defined in Section 28 of the IEEE 802.3 standard. The technologies currently supported by auto-negotiation are 10Base-T half duplex, 10Base-T full duplex, 100Base-TX half duplex, 100Base-TX full duplex, and 100Base-T4. The foundation for all of auto-negotiation's functionality is the fast link pulse (FLP) burst, which is simply a sequence of 10Base-T normal link pulses, also known as link test pulses (in 10Base-T technology) that come together to form a "word" (link code word) that identifies supported operational modes. On power-on or command from a management entity, a device capable of auto-negotiation issues an FLP burst.

Each FLP is composed of 33 pulse positions, with the 17 odd-numbered positions corresponding to clock pulses and the 16 even-numbered positions corresponding to data pulses. The time between pulse positions is 62.5 +/– 7 microseconds, and therefore 125 +/– 14 ps between each clock pulse. All clock positions are required to contain a link pulse. On the other hand, if a link pulse is present in a data position, it is representative of a logic 1, whereas the lack of a link pulse is representative of a logic 0. The amount of time between FLP bursts is 16 +/– 8 ms, which corresponds to the time between consecutive link test pulses produced by a 10Base-T device to allow interoperation with fixed-speed 10BASE-T devices. The 16 data positions in an FLP burst come together to form a link code word as shown in Fig. 22.3.

The 5-bit selector field contains 32 possible combinations, only 2 of which are currently defined and allowed to be sent. The next 8 bits, which make up the technology ability field, are used by a device to advertise its abilities to support various IEEE 802.3 technologies. The abilities are advertised in parallel; that is,

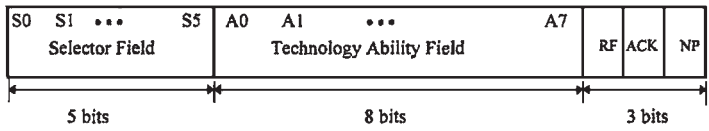


Figure 22.3 Link code word.

a single selector field value will advertise all the supported technologies. A logic 1 in any of these positions symbolizes that the device is capable of a particular IEEE 802.3 technology.

The device should advertise only the technologies that it supports. The devices at the two ends of the link segment may have an ability to support multiple technologies; the highest common denominator ability is always chosen, i.e., the technology with the highest priority that both sides can support. The relative priorities of the technologies supported by the EEE 802.3 selector field value are as follows (from highest to lowest):

1. 100BASE-TX full duplex
2. 100BASE-T4
3. 100BASE-TX
4. 10BASE-T full duplex
5. 10BASE-T

Remote fault: This bit can be set to inform a station that a remote fault has occurred.

Acknowledge: This bit is set to confirm the receipt of at least three complete, consecutive, and consistent FLP bursts from a station. This functionality will be discussed in detail later.

Next page: This is a means by which devices can transmit additional information beyond their link code words. This bit simply indicates whether or not more Next Pages are to come. When set to logic 0, it indicates that additional pages will follow, whereas a logic 1 indicates that there are no remaining pages.

22.2. ETHERNET PHYSICAL LAYER

Ethernet uses various types of copper links for relatively short-distance connections, and either multimode or single-mode optical fiber for longer distances. (Many combinations of data rate and media type have been standardized; see the Appendix for a complete list.) Copper versions of the physical layer include shielded twisted pair, unshielded twisted pair, and various forms of coax cables such as cat 5 or cat 6 cabling. Many types of repeaters, hubs, switches, and fanouts can be used to construct a range of different network topologies, with physical stars being among the earliest and most common. As networks grew larger, it became desirable to partition the collision domains into smaller subnetworks, as well as to overcome limits on the number of computers that could be connected to a given network segment. Bridges or switches were developed to forward traffic between network domains. While most communications are half duplex, it is possible to connect a single device to a switch port and achieve full duplex transmission. This doubles the aggregate link bandwidth and removes some of the

distance restrictions for collision detection on a link segment, which can be significant for fiber optic implementations. However, performance will not necessarily double in this situation, since traffic patterns are rarely symmetrical in practice.

Ethernet standards employ a naming convention that incorporates the transmission speed and maximum link segment length. Historically, the earliest Ethernet networks used interfaces such as 10Base5, where the 10 refers to a transmission speed of 10 Mbit/s, the word "Base" stands for baseband signaling (as opposed to broadband signaling; baseband means there is no frequency-division multiplexing or frequency-shifting modulation used), and the 5 refers to the maximum link length of 500 meters, achieved over a stiff (0.375-inch diameter) copper cable. Subsequent standards such as 10Base2 employed thinner copper coaxial cables with BNC connectors over a reduced maximum distance of 200 meters. Later standards continued to use the naming convention with a number to indicate the data rate in Mbit/s, followed by the word "Base," but changed to adopting a suffix that indicated the type of transmission media. For example, 10Base-T designates the use of twisted pair copper cable (in which two conductors are wound together to reduce crosstalk and electromagnetic interference from outside sources). The use of twin-axial copper cable or twinax (a balanced twisted pair of conductors with a cylindrical shield) is denoted as 10Base-CX. When there are several standards for a given transmission speed, their names may be distinguished by adding a letter or digit at the end of the name. For example, a 100-Mbit/s copper link using four twisted pairs of wire (at least cat 3 or higher) is known as 100Base-T4. While all of the link designations may refer to either half or full duplex operation, not all variants of the standard are commonly implemented; in our last example, 100Base-T4 was used exclusively as a half duplex link. Historically, there have also been wireline and wireless versions of Ethernet over radio frequency (RF) links. However, the currently recommended RF wireless standards, IEEE 802.11 and 802.16, do not use either the Ethernet link layer header or standard Ethernet control and management packets.

The collective group of standards operating at 100 Mbit/s (such as 100Base-T) are commonly known as Fast Ethernet, while the group operating at 1000 Mbit/s (such as 1000Base-T) are collectively known as Gigabit Ethernet. The notation changes slightly at the next highest data rate; for instance, 10 Gigabit Ethernet is denoted as 10GBase-T. With the widespread adoption of Gigabit Ethernet, half duplex links and repeaters became increasing less common and were discontinued altogether in the 10 Gigabit Ethernet standards. The CSMA/CD protocol was replaced by a system of full duplex links connected through Ethernet switches. Following the tradition of increasing each successive generation's data rate by a factor of 10, development has recently begun on a version of 100 Gbit/s Ethernet (100GBase-X). As this chapter goes to press, many of the details in this new standard are not yet finalized. For example, Gigabit Ethernet employs a DC

balanced 8B/10B encoding scheme (with NRZ signaling), which increases the overhead and thus the line rate from 1000 Mbit/s to 1250 Mbit/s. Higher data rates are considering adoption of a different signal encoding scheme such as 64B/66B, which would require special encapsulation or adaptation to interoperate with lower data rates.

Among the most common variants of copper links are 10Base-T, 100Base-TX, and 1000Base-T, which utilize twisted pair cables and a modular 8P8C electrical connector similar to the RJ-45. As we might expect, the maximum achievable distances have grown steadily shorter with higher data rates, and the standard has become increasingly selective about the type of copper cable required. A list of copper cable variants for Ethernet and other standards, as defined by ISO/IEC 11801 is given in Table 22.1. The most recent version, cat 7, contains four twisted pairs of copper wire terminated with either RJ-45 style compatible GG45 electrical connectors or with TERA connectors; this version was developed to allow transmission of 10 Gigabit Ethernet over 100 meters (the specially shielded cat 7 cable is rated for transmission frequencies up to 600 MHz). These cables are typically wired according to the EIA/TIA 568 A or B standards. Similar cable standards have been adopted for InfiniBand and other high-speed network

Table 22.1

Unshielded and Shielded Twisted Pair Cabling Standards.

- **Cat 1**: Currently unrecognized by TIA/EIA. Previously used for telephone communications, ISDN, and doorbell wiring.
- **Cat 2**: Currently unrecognized by TIA/EIA. Previously used frequently on 4 Mbit/s token ring networks.
- **Cat 3**: Currently defined in TIA/EIA-568-B, used for data networks using frequencies up to 16 MHz. Historically popular for 10 Mbit/s Ethernet networks.
- **Cat 4**: Currently unrecognized by TIA/EIA. Provided performance of up to 20 MHz, and frequently used on 16 Mbit/s token ring networks.
- **Cat 5**: Currently unrecognized by TIA/EIA. Provided performance of up to 100 MHz and frequently used on 100 Mbit/s Ethernet networks. May be unsuitable for 1000BASE-T gigabit ethernet.
- **Cat 5e**: Currently defined in TIA/EIA-568-B. Provides performance of up to 100 MHz and is frequently used for both 100 Mbit/s and gigabit Ethernet networks.
- **Cat 6**: Currently defined in TIA/EIA-568-B. Provides performance of up to 250 MHz, more than double category 5 and 5e.
- **Cat 6a**: Future specification for 10 Gbit/s applications.
- **Cat 7**: An informal name applied to ISO/IEC 11801 Class F cabling. This standard specifies four individually shielded pairs (STP) inside an overall shield. Designed for transmission at frequencies up to 600 MHz.

protocols. However, this cable has increased susceptibility to tight bend radius and other improper handling, and currently remains quite expensive. As a result, many installers avoid using the higher performance copper cables except as required for short connections to servers. For achieving longer distances in a cost-effective manner, optical fiber is the preferred physical layer media for Ethernet, and fiber-optic standards have traditionally been released prior to copper standards for higher data rates.

22.2.1. Fast Ethernet

Fast Ethernet [1–3], or 100BASE-T, is a 100-Mbps networking technology based on the IEEE 802.3 standard. It uses the same media access control (MAC) protocol, the carrier sense multiple access collision detection (CSMA/CD) method (running 10 times faster) that is used in the existing Ethernet networks (International Standards Organization/International Electrotechnical Community (ISO/lEC) 8802-3) such as 10BASE-T, connected through a media-independent interface (MII) to the physical layer device (PHY) running at 100 Mb/s. The supported PHY sublayers are 100BASE-T4, l00BASE-TX, and 100BASE-FX. The general name of the PHY group is 100BASE-T. One of the most important advantages of Fast Ethernet is that it uses the standard 802.3 MAC. This allows data to be interchanged between a 10BASE-T and 100BASE-T nodes without protocol translation, thereby allowing the possibility of a phased introduction of 100BASE-T in the existing 1OBASE-T networks and a cost-effective migration path that maximizes use of existing cabling and network management systems. Auto-negotiation is another key element of the 100BASE-T Fast Ethernet IEEE 802.3u standard. Auto-negotiation allows an adapter, hub, or switch capable of data transfer at both 10 and 100 Mb/s rates to automatically use the fastest rate supported by the device at the other end. Auto-negotiation signals the capabilities it has available, detects the technology that exist in the device it is being connected to, and automatically configures to the highest common performance mode of operation.

22.2.2. 100BASE-T4

This physical layer defines the specification for 100BASE-T Ethernet over four pairs of category 3, 4, or 5 unshielded twisted pair (UTP) wire. This is aimed at those users who want to retain the use of voice-grade twisted pair cable. In addition, it does not transmit a continuous signal between packets, which makes it useful in battery-powered applications.

With this signaling method, one pair is used for carrier detection and collision detection in each direction, and the other two are bidirectional. This allows for a half duplex communication using three pairs for data transmission. The unidirectional pairs are the same ones used in 10BASE-T (it uses only two pairs) for

consistency and interoperability. Because three pairs are used for transmission, to achieve an overall 100 Mb/s each pair needs only to transmit at 33.33 Mb/s. If Manchester encoding was to be used at the physical layer, as in 10BASE-T, the 30-MHz limit for the category 3 (cat 3) cable would be exceeded. To reduce the rate, a 8B6B block code is used that converts a block of 8 bits into six ternary symbols, which are then transmitted out to three independent channels (pairs). The effective data rate per pair thus becomes $(6/8) \times 33.33 = 25$ MHz, which is well within the cat 3 specifications of 30 MHz.

22.2.3. 100BASE-X

Two physical layer implementations, 100BASE-TX and 100BASE-FX, are collectively called 100BASE-X when referring to issues common to both. 100BASE-X uses the physical layer standard of FDDI by using its physical media-dependent sublayer (PMD) and medium-dependent interfaces (MDI). The 125 Mb/s full duplex signaling system for a twisted pair, defined in FDDI, forms the basis for 100BASE-TX, and the system defined for transmission on optical fiber forms the basis for 100BASE-FX. Basically, 100BASE-X maps the characteristics of FDDI PMD and MDI to the services expected by the CSMA/CD MAC. It defines a full duplex signaling standard of 125 Mb/s for multimode fiber, shielded twisted pair, and unshielded twisted pair wiring. The physical sublayer maps 4 bits from MI1 into 5-bit code blocks and vice versa using the 4B/5B encoding scheme (the same as FDDI).

22.2.4. 100BASE-TX

This physical layer defines the specifications for 100BASE-T Ethernet over two pairs of category 5 UTP wire or two pairs of shielded twisted pair (STP) wire. With one pair for transmit and the other for receive, the wiring scheme is identical to that used for 10BASE-T Ethernet.

22.2.5. 100BASE-FX

This physical layer defines the specification for 100BASE-T Ethernet over two strands of multimode (62.5/125 pm) fiber cable. One strand is used for transmit, whereas the other is used for receive.

100BASE-T Fast Ethernet and 10BASE-T Ethernet differ in their topology rules. 100BASE-T preserves 10BASE-T's 100-m maximum UTP cable runs from hub to desktop. The basic rules revolve around two factors: the network diameter and the class of the repeater (or hub). The network diameter is defined as the distance between two end stations in the same collision domain. Fast Ethernet specifications limit the network diameter to approximately 205 m (using UTP cabling), whereas traditional Ethernet could have a diameter up to 500 m.

Repeaters are the means used to connect segments of a network medium together in a single collision domain. Different physical signaling systems can be joined into a common collision domain using a repeater. Bridges can also be used to connect different signaling systems; however, each system connected to the bridge will have a different collision domain. In traditional Ethernet, all hubs are considered to be functionally identical. In Fast Ethernet, however, there are two classes of repeaters: class I and class 11. Class I repeaters perform translations when transmitting, enabling different types of physical signaling systems to be connected to the same collision domain. Because they have internal delays, only one class I repeater can be used within a single collision domain when maximum cable lengths are used; that is, this type of repeater cannot be cascaded, and the maximum network diameter is 200 m. On the other hand, class II repeaters simply repeat the incoming signals with no translations; that is, they provide ports for only one physical signaling system type. Class II repeaters have smaller internal delays and can be cascaded using a 5-m cable with a maximum of two repeaters in a single-collision domain if maximum cable lengths are used. Cable lengths can always be sacrificed to get additional repeaters in a collision domain.

With traditional 10BASE-T Ethernet, networks are designed using three basic guidelines: maximum UTP cable runs of 100 m, four repeaters in a single collision domain, and a maximum network diameter of 500 m. With 100BASE-T Fast Ethernet, the maximum UTP cable length remains 100 m. However, the repeater count drops to two, and the network diameter drops to 205 m. Using optical fiber, the maximum collision domain diameter for a point-to-point link is 412 meter, 272 meters for one class 1 repeater, 320 meters for one class II repeater, and 228 meters for two class II repeaters. On the surface, the 100BASE-T Fast Ethernet rules seem restrictive, but with use of repeaters, bridges, and switches, Fast Ethernet can be easily implemented in a network.

22.3. GIGABIT AND 10 GIGABIT ETHERNET

The purpose of Gigabit Ethernet (IEEE 802.32) is to extend the 802.3 protocol to an operating speed of 1000 Mb/**s** in order to provide a significant increase in bandwidth while maintaining maximum compatibility with the installed base of 10/100 CSMA/CD nodes, research and development, network operation, and management [4–6]. This standard was designed to use the 802.3 Ethernet frame format, including preserving minimum and maximum frame sizes, and to comply with 802 FR (with the possible exception of Hamming distance). Forwarding between the 1000, 100, and 10 Mbit/s data rates is supported at both full and half duplex operation (in practice, half duplex is very seldom used). Link distances and supported media are described in Appendix D.

22.3.1. 10 Gigabit Ethernet

Physical Media Dependent sublayers, or PMDs, were defined by the IEEE 802.3ae task force in 2002. These sublayers define physical layer specifications for 10 Gigabit Ethernet transmissions and are given in Appendix D. These include quite a few new technologies, and since this protocol is in the early stages of development, the preferred direction is not clear as of this writing. For example, the standard has proposed long reach modules (LRM), including single-mode fiber versions in the 1550-nm band operating up to 40 km, and single-mode fiber in the 1300-nm band for more limited distances (up to 10 km unrepeated). Transceiver form factor options include the enhanced version of small form factor pluggable (SFP) transceivers known as SFP+, which has found applications in many 10 gigabit switches and routers. There are also multimode fiber versions defined for shorter distances (26–82 m over installed lower bandwidth fiber, or up to 300 m on newer OM3 fiber). A coarse wavelength-division multiplexing (CWDM) transceiver using 4 laser wavelengths is defined, where each wavelength operates at 3.125 Gbit/s; this supports distances up to 300 m on multimode fiber or 10 km on single-mode fiber. We also note that under certain conditions it is possible for a single-mode transceiver to operate over multimode fiber at restricted distances (around 100–200 m) if a special mode conditioning patch cable (MCP) is attached at both ends of a duplex link (for a discussion of MCPs see Chapter 4).

22.3.2. 40 and 100 Gigabit Ethernet

In late November 2006, an IEEE study group agreed to target 100Gbps Ethernet as the next version of the technology. While this standard is not yet finalized as this chapter goes to press, the IEEE 802.3 Higher Speed Study Group (HSSG) has adopted several objectives that direct their current work. These include 100GbE optical fiber Ethernet standards for at least 100 meters and at least 10 kilometers, full duplex operation only, and using current frame format and size standards. It is expected that the Study Group will present a Project Authorization Request (PAR) to the 802 Standards Executive Committee in the near future. The objectives for this work include support for 100 Gb/s at the MAC interface over at least 100 m on OM3 multimode fiber, 10 meters over copper links, and 10–40 km on single-mode fiber (full duplex operation only). An objective of this work is to meet the requirements of an optical transport network (OTN). It is expected that the IEEE 802.3 frame format will be preserved at the MAC interface, as well as the current maximum and minimum frame sizes. The existing industry standard bit error rate of 10e-12 at the MAC/PLS service interface is also likely to remain unchanged. The study group has also agreed to adopt a 40 Gbit/s data rate support at the MAC layer, meeting the same conditions as 100 Gbit/s Ethernet with the exception of not supporting 10–40 km distances. The proposed 40 Gbit/s standard will also allow operation over up to 1 meter on backplanes.

Table 22.2
Ethernet Data Rates over Time.

Product Name	Throughput* (MBps)	Line Rate (GBaud)	IEEE Specification Completed (Year)	Market Availability (Year)
Ethernet	2.5	0.010	1985	1980
Fast Ethernet	25	0.10	1995	1996
Gigabit Ethernet	250	1.25	1998	1999
10 Gigabit Ethernet	2,500	10.3	2002	2006
40 Gigabit Ethernet	10,000	4×10.3	2010	Market demand
100 Gigabit Ethernet	25,000	10×10.3	2010	Market demand

*The throughput of the links includes data communications in both directions.

22.3.3. Ethernet Roadmaps

Table 22.2 shows how the Ethernet roadmap has traditionally increased its data rate by factors of 10 every few years (a possible exception being the recent interest in 40 Gbit/s Ethernet). These order-of-magnitude jumps in Ethernet contrast to the doubling of speed seen in Fibre Channel's Base-2 speeds.

Some analysts contend that these large incremental increases in data rate lead to difficulty in high volume, low-cost manufacturing of equipment at rates exceeding 10 Gbit/s, and longer lead times to widespread adoption (perhaps several years after the standard is complete). For example, about five years after the 10 Gigabit Ethernet standard was adopted, efforts to reduce product cost continue and deployment is limited to applications with a critical need for the higher data rate. Similar difficulties may be encountered for higher data rates in the future. Concern with the backward compatibility of higher data rates has further affected adoption of these link rates.

22.4. METRO AND CARRIER ETHERNET

With the wide adoption of Ethernet interfaces in the LAN and local loop access network, some companies have begun working toward the goal of extending these links further into the metropolitan area network (MAN) and wide area network (WAN). This would allow Ethernet-based links to extend hundreds or even thousands of kilometers, and provide an alternative for SONET/SDH links or a way for commercial users to connect their branch offices to their intranets. The potential advantages of this approach include providing high-bandwidth interfaces with a finer granularity than SONET/SDH, the potential convergence of voice over IP with data networking, increased flexibility in designing topologies or in attachment to existing corporate networks, and the reuse of existing Ethernet skill sets already present at data center users. On the other hand, classic Ethernet does not offer the full functionality of a SONET/SDH interface, including differentiated

quality and class of service (which facilitates the creation of service-level agreements between customers and network providers), inherent fast protection switching, and other high reliability features. This has created interest in developing a version of Ethernet capable of emulating the important functions of the MAN and WAN environment, typically known as carrier class Ethernet. The creation of industrywide specifications for carrier class Ethernet is administered through an industry consortium, Metro Ethernet Forum (MEF).

The MEF was formed in 2001 to accelerate the worldwide adoption of Ethernet services, in a form that is hardened to withstand the same rigors as the current telecommunications environment (so-called carrier class Ethernet). There have been regular updates to the carrier Ethernet standard over the years, including the recent 10.1 release in mid-2006 (a list of MEF standards is given in Appendix E). The goal is to retain most of the simplicity and cost model of Ethernet while expanding its role in access and metro networks, interconnection of enterprise LANs, and convergence of business, residential, and wireless network infrastructure. The resulting specification for carrier class Ethernet is based on several attributes. First, they attempt to create standardized services that do not require changes to existing network equipment, and they can accommodate time-sensitive TDM traffic (this is part of the "triple play" convergence of voice, video, and data traffic). Carrier Ethernet is also intended to be scalable over global distances, in useful bandwidth increments up to 10 Gbit/s. Rapid network recovery and reconvergence is also incorporated, including a 50-ms switch time similar to SONET/SDH. Carrier Ethernet provides differentiated levels of service, useful in creating end-to-end service-level agreements (SLAs) with carriers based on attributes such as error rate, frame loss, delay, and delay variation. Finally, Carrier Ethernet is intended to provide rapid provisioning and carrier-class operations management (OAM). Equipment and services that comply with the MEF standards can earn a compliance certification mark and can be used in Metro Ethernet networks (MENs). The MEF currently has over 110 service provider and carrier members; some of the largest commercial Metro Ethernet deployments so far include over 1 million homes on a major network in Hong Kong.

A typical service provider Metro Ethernet network is a collection of layer 2 or 3 switches or routers connected through optical fiber. The topology could be a ring, hub-and-spoke (star), full mesh or partial mesh. The network will also have a hierarchy: core, distribution, and access. The core in most cases is an existing IP/MPLS backbone. Ethernet on the MAN can be used as pure Ethernet, Ethernet over SDH, Ethernet over MPLS, or Ethernet over DWDM. Pure Ethernet-based deployments are least expensive but less reliable and scalable, and thus are usually limited to small-scale deployments. SDH-based deployments are useful when an existing SDH infrastructure is already in place, its main shortcoming being the loss of flexibility in bandwidth management due to the rigid hierarchy imposed by the SDH network. MPLS-based deployments are costly but highly reliable and

scalable, and they are typically used by large service providers. In typical approaches, connections between two user network interfaces (UNIs) form an Ethernet Virtual Connection (EVC). These can be either point-to-point or multipoint solutions. The two types of services are known as E-Line (also called Point-to-Point, Virtual Leased Line, Virtual Private Wire Service (VPWS), or Ethernet over MPLS (EoMPLS)) and E-LAN (also known as Transparent LAN Services, Multipoint-to-Multipoint, and Virtual Private LAN Services).

Ethernet can be deployed in the MAN in several ways. For example, we may consider a pure Ethernet MAN uses only layer 2 switches for all its internal structure. This allows for a very simple design, and also for a relatively simple initial configuration. The original Ethernet technology was not well suited for service provider applications; as a shared-media network, it was impossible to keep traffic isolated, which made implementation of private circuits impossible. Ethernet MANs became feasible in the late 1990s due to the development of new techniques to allow transparent tunneling of traffic through the use of Virtual LANs as either point-to-point or multipoint-to-multipoint circuits. Combined with new features such as VLAN Stacking (also known as VLAN Tunneling), and VLAN Translation, it became possible to isolate the customer's traffic from each other and from the core network internal signaling traffic. For small-scale deployments, a pure Ethernet MAN can be quite effective.

There are several main drawbacks to a pure Ethernet MAN approach. First, by design, layer 2 switches use fixed tables to direct traffic based on the MAC address of the end points. As the network gets larger, the number of MAC addresses transiting through the network may grow beyond the capacity of the core switches. If the core routing table gets full, the result is a catastrophic loss of performance due to the flooding of packets over the entire network structure. Second, network stability is relatively fragile, especially if compared to the more advanced SDH and MPLS approaches. The recovery time for the standard spanning tree protocol is in the range of tens of seconds, much higher than what can be obtained in alternative networks (usually a fraction of second). There are a number of optimizations, some standardized through the IEEE, and others vendor-specific, that seek to alleviate this problem. The clever use of such features allows the network to achieve good stability and resilience, at the cost of a more complex configuration and possible use of nonstandard, vendor-specific, mechanisms. Finally, for the pure Ethernet MAN, traffic engineering is very limited. There are few tools to manage the topology of the network; also, the fact that forwarding is done hop-by-hop, added to the possibility of broadcasts even for unicast packets (for instance, while learning new addresses), makes predicting the real traffic pattern very difficult. There are techniques that allow for some control of the preferential traffic paths; these techniques rely on the use of multiple spanning trees, or "per VLAN spanning trees," and are closely connected to the solutions used to achieve better stability and resiliency in the network.

An alternative that forms an intermediate stage in the transition from a traditional, time-division-based network such as SONET/SDH, to a pure Ethernet network involves running an SDH-based Ethernet MAN. In this model, the existing SDH infrastructure is used to transport high-speed Ethernet connections. The main advantage of this approach is the high level of reliability, achieved through the use of the native 50-ms SDH protection switching mechanisms (see Chapters 15 and 19). Hybrid designs use conventional Ethernet switches at the edge of the core SDH ring to alleviate some traffic management and network topology constraints.

Finally, an MPLS-based Metro Ethernet network uses MPLS in the service provider network. The subscriber will get an Ethernet interface on copper (100Base-TX) or fiber (100Base-FX). The customer's Ethernet packet is transported over MPLS, and the service provider network uses Ethernet again as the underlying technology to transport MPLS. This approach is more scalable than pure Ethernet MANs, which are limited to a maximum of 4096 VLANs in a network (in an MPLS implementation, the VLANs have local meaning only). This approach can also take advantage of MPLS local protection and restoration mechanisms such as Fast ReRoute (FRR), which offer recovery times on the order of SDH (50 ms), as opposed to pure Ethernet networks that rely on STP or RSTP protocols that can take several seconds or more to converge. An MPLS-based Metro Ethernet can backhaul not only IP/Ethernet traffic but virtually any type of traffic coming from customers networks or other access networks.

22.5. E-PONS AND ETHERNET FIRST MILE

Various efforts have been made to use Ethernet to address the bandwidth bottleneck between residential consumers and the carrier access network (the so-called last mile, also referred to in the literature as the subscriber access network, the local loop, or paradoxically the "first mile"). These include proposals such as Fiber to the Home (FTTH), to the curb (FTTC), to the pole (FTTP), or similar access points collectively known as FTTX. One recent proposal was led by the industry consortium Ethernet First Mile Alliance (EFMA), which proposed a solution known as Ethernet in the First Mile (EFM) that was adopted as IEEE 802.3ah. In 2004, the EFMA became part of the MEF. Three basic standards are defined for copper, fiber, and passive optical networks, including variants that use two fibers and others that use a single fiber for bi-directional transmission with two wavelengths near 1550 nm and 1300 nm (see Appendix D).

Another variant of this approach is Ethernet over a Passive Optical Network (E-PON). This is one of a class of passive optical networks (see Chapter 16) proposed for residential broadband access. An E-PON includes connections between an optical line terminal (OLT) in the carrier central office or point-of-presence (typically an Ethernet switch or converter) and an optical network unit (ONU) near the customer premise. The E-PON is configured in full duplex mode

(without CSMA/CD) using single fiber point-to-multipoint (P2MP). In other words, E-PON uses one optical fiber and two multiplexed wavelengths to minimize the number of fibers and requires N + 1 transceivers to access N network nodes. The subscriber, or ONU, transmit uplink operates at a wavelength of 1310 nm, while the downlink operates at 1490 nm. ONUs can only communicate peer-to-peer through the OLT, which allows one subscriber at a time to transmit using a time-division multiple access (TDMA) protocol. There are no optical amplifiers or other network equipment in a PON, only passive optical splitters. Variants for 10-km and 20-km distances are available, each supporting a split ratio of 16 : 1.

22.6. ENHANCED ETHERNET

Given the pervasiveness and relatively low cost of Ethernet as the predominant LAN technology, as well as its recent acceptance for certain WAN applications, some have proposed that in the future Ethernet could serve as the long-sought convergence fabric for voice, video, and data communication. While this has been promised in the past by various other protocols (including ATM, Fibre Channel, and more recently InfiniBand), there are some aspects of Ethernet that make it well suited to this role. In order to achieve this vision, it is necessary for future Ethernet standards to take on the attributes of the various protocols they will replace, much as carrier class Ethernet has found it necessary to emulate some behaviors of traditional SONET/SDH networks. It would also be desirable to faciliate encapsulation of other frame protocols in an Ethernet digital wrapper and to provide accelerated performance and lower latency for these new protocols. Proposals along these lines were made to the IEEE standards committee as early as 2005; however, they have recently gained new market momentum. While there is not yet an IEEE standard supporting this work, various prestandard activities have begun in the industry in an effort to drive ad hoc adoption of these goals without going through the lengthy, formal IEEE standards body review process. This work is not fully defined as this chapter goes to press, but is known collectively as Enhanced Ethernet, or sometimes by the Cisco trademarked name Data Center Ethernet.

In its most general form, Enhanced Ethernet is an emerging industry standard that proposes modifications to existing Ethernet networks, in an effort to position Ethernet as the preferred convergence fabric for all types of data center traffic. While some features may be added or removed before Enhanced Ethernet products reach the market or are standardized by the IEEE, at this time several major changes have been proposed for Enhanced Ethernet. These changes would not affect TCP/IP, which runs on top of Enhanced Ethernet for some applications; it is hoped that they will offer a much simpler, low-cost approach that does not require offload processing or accelerators (in contrast to existing TCP/IP Offload Engines, for example). Suggested enhancements to network acceleration and

application programming interfaces (APIs) to streamline performance have led to the informal use of names such as Low Latency Ethernet in reference to these future standards. Proposed features of Enhanced Ethernet include the following:

- The addition of credit-based flow control at the physical layer, similar to the function provided by Fibre Channel and FICON links today. Fibre Channel is quite sensitive to data packets arriving out of order and employs a buffer-to-buffer scheme to manage the flow of data through the network (see Chapter 20). Future versions of Ethernet may incorporate a similar scheme, as a way to emulate the functions of an enterprise-class storage area network (SAN).
- The addition of congestion detection and data rate throttling features. For high performance in the data center, Enhanced Ethernet is intended to run without TCP/IP, and thus has no inherent way to perform link error recovery. Modifications to the link layer protocols, as yet undefined, would be intended to enable lossless error recovery (as opposed to the best effort delivery and frame discarding functions of TCP/IP), possibly implemented as low as the physical link layer.
- The addition of virtual lanes with quality of service differentiation. Potentially up to eight service classes have been proposed, which would emulate behaviors in either ATM/SONET or Fibre Channel link designs. It is important to note that these functions do not affect TCP/IP, which is above the DCE level; this should offer a much simpler, low-cost approach that does not require offload processing or accelerators.

It is generally felt that any future version of Ethernet would be based on at least a 10 Gbit/s interface. While these interfaces remain fairly expensive today, the high volumes created by a convergent network would presumably drive down costs over time. As Enhanced Ethernet moves into new application areas, such as disaster recovery, the need for technologies capable of extending links over tens to hundreds of kilometers begins to emerge. Once Enhanced Ethernet is standardized, this function will presumably be incorporated into switches, routers, channel extenders or wavelength multiplexers. The need to drive unrepeated links up to 10 km or more is established in other datacom standards, as well as in earlier versions of Ethernet. In order to maintain this distance, a 10 Gbit/s link might have to incorporate a different data encoding scheme, such as 64B/65B with scrambling rather than the more commonly used 8B10B schemes used today at lower data rates.

REFERENCES

1. IEEE Std 802.3. Local and metropolitan area networks, supplement-Media access control (MAC) parameters, physical layer, medium attachment and repeater for 100 Mb/s operation, type 100BASE-T

(clauses 21-30). Copyright 1995 by the Institute of Electrical and Electronics Engineers, Inc. The IEEE disclaims any responsibility or liability resulting from the placement and use in the described manner. Information is reprinted with the permission of the IEEE.
2. Molle, M., and G. Watson. 1996. 100Base-ThEEE 802.12/packet switching. *IEEE Commun. Mag.* 34(8):64–73.
3. E. Halsall. 1995. *Data communications, computer networks and open systems*. 4th ed. Reading, Mass.: Addison-Wesley.
4. IEEE 802.32, http://stdbbs.ieee.org/pub/802main/802.3/gigabit.
5. IEEE Std 802.12. Demand-priority access method, physical layer and repeater specifications for l00Mb/s operation. Institute of Electrical and Electronic Engineers, 345 East 47th Street, New York, N.Y. 10017.
6. Watson, G., A. Albrecht, J. Curcio, D. Dove, S. Goody, J. Grinham, M. P. Spratt, and P. A. Thaler. 1995. The demand priority MAC protocol. *IEEE Network* 9(1):28–34.

GENERAL REFERENCES

- Barbieri, Alessandro. "10 GbE and Its X Factors." *Packet: Cisco Systems Users Magazine*, Third Quarter 2005, Vol. 17, No. 3:25–28.
- Halabi, Sam (2003). *Metro Ethernet.* Cisco Press (ISBN 1-58705-096-X).
- The Metro Ethernet Forum, http://www.metroethernetforum.org/
- Ethernet WAN service offering example from Alcatel/Lucent: http://www1.alcatel-lucent.com/bnd/vpls/

This page intentionally left blank

Case Study Unbundling the Local Loop for Triple Play Networks

Courtesy of Nortel Optical Networks

Application: Design core and edge network infrastructure to support next-generation broadband services for a telecommunications carrier supporting over 500 exchanges in 300 cities spread across 7000 km.

Description: Traditional competitive barriers between telecommunication companies and other service providers are breaking down as a result of recent regulatory legislation and the availability of many lower cost fiber-optic communication technology options. Unbundling the local loop or "last mile" has created an opportunity for service providers to offer a range of new broadband services, including so-called triple play (simultaneous delivery of voice, video, and data over an IP network). This is enabled by a variety of new offerings such as voice over IP, IPTV, and video on demand. While this new competitive environment has created business opportunities, it has also placed unusual stress on existing networks. Unexpectedly high traffic volumes, unpredictable traffic patterns with high variability, instabilities in the network, consumer demand for consistently reliable network connections, and inadequate support staff all contribute to the design pressures on new triple-play infrastructures.

A major telecom service provider whose last mile was earmarked for unbundling needed to deploy a new fiber-based network to aggregate IP DSLAMs (digital subscriber line access multiplexers). This was part of a strategy to offer higher data rates in the local loop (up to 24 Mbit/s across ADSL 2+). Meanwhile, the carrier needed to retain and protect its significant investment in SONET/SDH transport and switching for the core network, but also wanted to aggregate Gigabit Ethernet traffic from the access network and deploy Optical Ethernet to backhaul

DSL traffic. This solution was implemented using a metropolitan area wavelength multiplexing platform (the Nortel Optera Metro 5000) to provide 32 wavelengths at up to 10 Gbit/s/wavelength in the core, thus increasing the utilization of the installed fiber and the bandwidth efficiency of the network. For the Access nodes, coarse WDM was used to accommodate the lower traffic volumes in a cost-effective manner; CWDM also provides a smaller footprint and lower power consumption for the network equipment. In this case, the coarse and dense WDM products make use of a common hardware design that simplifies installation and reduces some of the equipment sparing costs. The WDM network was designed to support interconnected rings, so that a protocol agnostic layer would overlay the existing SONET/SDH backbone. Link budgets were designed to allow point-to-point connections (up to 90 km in most cases) without the additional cost and complexity of electrical regeneration or optical amplifiers. The edge network was further supported by a series of layer 2 switches (initial deployments used the Optical Multiservice Edge 6500, although the design allows for scalable, lower bandwidth deployments in the future with products such as the Edge 6110 or 1000). The WDM networks were configured with 50-ms protection switching to equal the reliability of the SONET/SDH equipment. The suppliers guaranteed network availability exceeding 99.999% by design, and in the first year of operation, this network recorded zero customer downtime as a result of network resilience features.

23

FDDI and Local Area Networks[1]

Rakesh Thapar
Marconi, Warrendale, Pennsylvania

23.1. INTRODUCTION

Local area networks (LANs) are commonly used in many applications. This chapter describes one of the earliest fiber-based LAN technologies, fiber distributed data interface (FDDI).

FDDI is an ANSI standard and has an IS0 approval. The basic FDDI standard [12–14] can be categorized as shown in Fig. 23.1. It is based on a dual-ring topology composed of two counterrotating rings that operate at a data rate of 100 Mb/s. It uses dual counterrotating rings to enhance reliability: the primary ring and the secondary ring. The secondary ring is used in case of failures. Multiple failures, however, result in multiple rings. Multimode fiber connects each station together, and the total ring can be 100 km in length. A maximum of 500 stations can be connected in the ring, and hence it forms an ideal backbone network.

FDDI consists of a set of stations serially connected by a transmission medium to form a closed loop. Information is transmitted as symbols (the equivalent of four information bits) from one active station to the next. Each station, including the destination, simply regenerates each symbol on the ring in the direction of rotation. The destination, however, makes a copy as the information passes it. The frame is removed by the station that originated the transmission.

23.1.1. Station Types

There are two types of FDDI stations: a dual-attach station (DAS), which is connected to both rings, and a single-attached station (SAS), which is attached

[1]This chapter is an edited version of the chapter that appeared in the second edition of the *Handbook of Fiber Optic Data Communication*.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

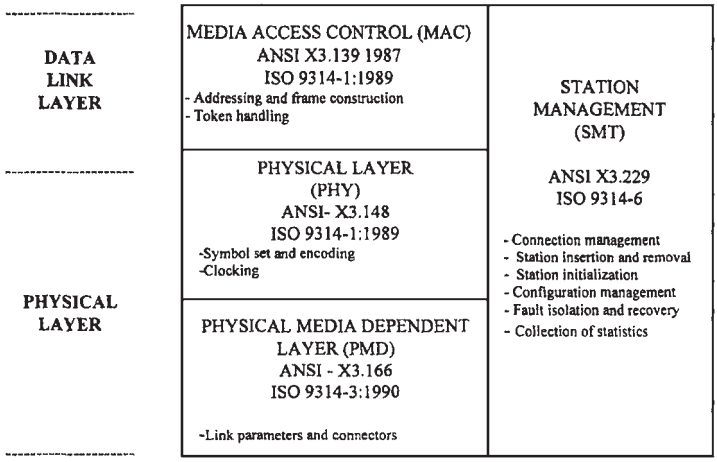


Figure 23.1 Physical layer of the FDDI standard.

only to the primary ring. A dual-attached station has at least two ports—an A port, where the primary ring comes in and the secondary ring goes out, and a B port, where the secondary ring comes in and the primary goes out. A station may also have a number of M ports, which are attachments for single-attached stations. Stations with at least one M port are called concentrators. SASs have S ports, which can be connected to the FDDI ring only through concentrators using M ports. A dual ring of DAS devices cannot be readily moved, changed, or added into. A single user disconnecting a dual-attached workstation causes a break in the ring. Dual-attach stations require twice the number of connectors and cables.

23.1.2. Physical Layer Specifications

FDDI supports two basic media types—copper and fiber. With copper it works with shielded twisted pair (STP) cable and allows data at 100 Mbs for up to 100 m. When using STP cabling, we get FDDI over copper commonly known as copper distributed data interface. Copper cables will be used to connect workstations to hub only. UTP has also been considered by the X3T9.5 committee. When using fiber, there are three alternatives: single-mode fiber, multimode fiber, and plastic or low-cost fiber. The single-mode fiber can be used to work at FDDI speed over a distance of 50 km but is expensive. Multimode fiber is the most commonly used media. Plastic or low-cost fibers were marketed for a while but, due to the nature of plastic, the distance was limited to 100 m and interest in them was short lived. An FDDI station attaches to the fiber-optic cable by a media interface connector (MIC) whose main function is to mechanically align the fibers. MIC plugs have

Table 23.1

FDDI Connector Types.

FDDI Connection Type	Key Type
Workstation to wall	S/S
Distribution panel to concentrator M port	M/M
Distribution panel to concentrator A port	A/A
Distribution panel to concentrator B port	B/B
Concentrator A to concentrator M port	A/M
Concentrator B to concentrator M port	B/A
Concentrator 1A to concentrator 2B port	A/B
Concentrator 1B to concentrator 2A port	B/A
Workstation to concentration M port	S/M

Table 23.2

FDDI I/O Interfaces.

	Output Interface		*Input Interface*	
	Min	Max	Min	Max
Center wavelength	1270 nm	1380 nm	1270 nm	1380 nm
Average power	–20.0 dBm	–14.0 dBm	–31.0 dBm	–14.0 dBm

mechanical latches that mate with the latch points in the MIC receptacle. As an option, the MIC receptacles of a station can be keyed to prevent improper attachment of input and output. Keys are used to determine the PMD entries contained in the FDDI node. Four keys are defined: type A or red keyflabel (primary idsecondary out) and type B or blue keyflabel (secondary idprimary out) are used to connect Dual Attach Stations (DAS) into an FDDI ring. Type M or green keyflabel (master) is used on the concentrator side to connect a SAS to a concentrator. Type S or no key/white label (slave) is used on the station side to connect a SAS to the concentrator (i.e., to connect to a port M). Table 23.1 summarizes the general MIC connector key usage, and Table 23.2 gives the FDDI interface power levels.

23.1.3. Coding and Symbol Set

FDDI uses 4B/5B encoding, i.e., each 4 bits are encoded into a 5-bit symbol. FDDI communicates all of its information using symbols. This 5-bit encoding provides 16 data symbols (normal hex range: 0-F), eight control symbols (Q, H, I, J, K, T, R, and s), and eight code violation symbols (***V***). After the data are 4B/5B encoded, they are further encoded using non-return-to-zero invert coding before the bits are let out on the media.

23.1.4. Line States

Line states are used by the physical layer to communicate information. A line state is a continuous stream of a certain symbol(s) sent by transmitter (physical layer) that, upon receipt of another station, will uniquely identify the state of the communication line. The following are the commonly encountered line states:

Quiet line state: A continuous stream of "Q' symbols. This line state is used during connection management signaling.

Halt line state: A continuous stream of "H' symbols. This line state is used to indicate a trace and also used during physical connection management (PCM) signaling.

Master line state: An alternating series of "H' and "Q' symbols. This is used to break a link and restart a connection.

Idle line state: A continuous stream of "I" symbols. This line state is used to separate information bits and to provide clock synchronization.

Active line state: This line state is asserted by a "JK' pair and continues to be asserted until it receives "R," "S," or "T" symbols. It is used by MAC layer to transmit data frames.

Link state unknown: This is an indication by the physical layer if the receiver is unable to determine the line state.

Noise line state: This is an indication by the physical layer if the receiver has received enough symbols but is unable to uniquely recognize the line state (i.e., the receiver is getting garbled information on the line).

23.2. MAC SPECIFICATIONS

FDDI is a controlled-access network. The access to transmission is through a token. The token is a particular unique combination of symbols that circulates on the ring in one direction from station to station. A station receives the token in its entirety on its input port and then regenerates the token on the output port to continue circulation. In order to transmit, a station waits for the token and on receipt of the token removes it from the ring, thereby capturing the token. It then puts out the frame(s) it wants to transmit and then at the end introduces a new token. A station can hold on to a token for a finite time defined as the token-holding time. The station that transmits a frame is also responsible for removing the frame from the ring. The basic protocol data units used by the FDDI MAC are the token and the frames.

23.2.1. Token

The general format of a token is shown in Fig. 23.2. The preamble (PA) is a minimum of 16 symbol of idles. Because of ring latency and timing constraints,

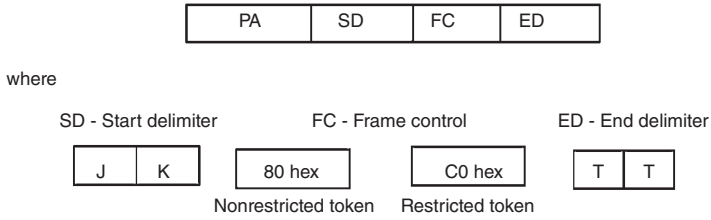


Figure 23.2 Token format.

Figure 23.3 Format for FDDI frame transmission.

subsequent stations may see a smaller length of preamble. Tokens are recognized when received with a preamble of zero or greater length. The start delimiter is made up of two symbols, J and K, which will not be seen anywhere except to mark the beginning of a new token or a frame. Similarly, the end delimiter is made of two symbols, which are both a T. For a token there are two possible values (as shown previously) for the frame control field, depending on whether it is a nonrestricted or restricted token. The nonrestricted mode is the normal mode of operation in which the bandwidth is shared equally among all requesters. The restricted mode is used to dedicate bandwidth to a single extended dialog between specific requesters (these are described later in this section).

23.2.2. Frame

FDDI defines three basic types of frames: (1) the MAC frames, which carry MAC control data; (2) logical link control frames (LLC), which carry user information; and (3) station management frames (SMT), which carry management information between stations and layers. MAC frames include frames used in ring initialization and beacon frames used in the process of ring fault isolation. These frames never leave the FDDI ring (i.e., they never go across a bridge or a router) because they are specific for a ring only. LCC frames carry user information between nodes on a network. These frames do cross bridges or routers. SMT frames, like MAC frames, do not cross bridges or routers because they are used to control the operation of a ring only.

Figure 23.3 shows the general format used to transmit the LLC, MAC, and SMT frames. The SMT frames are discussed later, and this section will focus on MAC and LLC frames: The preamble, start delimiter, and end delimiter are the

same as in a token. The frame control is used to distinguish between a token, MAC frame, LLC frame, or SMT frame. The Fibre Channel Standard (FCS) is a 32-bit polynomial checksum.

The frame status consists of three indicators that may have one or two values. The indicators can either be set (*S*) or reset (*R*). Every frame is originally transmitted with all the indicators set to *R*. The indicators can be set by intermediate stations when they retransmit the frame. The three indicators are error, address recognized, and copy. Error is set when a station determines that the frame is in error. This may be due to a format error or failure of cyclic redundancy check (CRC). The address recognized bit is set when a station receives the frame and determines that the destination address is its own MAC address (or if it was broadcast, then the first recipient sets it). Copy indicator is set when the destination station is able to copy the contents into its buffers.

23.3.3. Ring Operation

Access to the ring is controlled by passing a token around the ring. The token gives the downstream station (relative to the one passing the token) an opportunity to transmit a frame or a sequence of frames. If a station wants to transmit, it must first capture the token, that is, strip the token off the ring and then begin transmitting its eligible queued frames. After finishing transmission, the station immediately issues a new token for use by downstream stations, without waiting for frames to return on the ring. Optionally, only for various station management functions, the station may wait to see one or more or all the transmitted frames return before it issues the token. If a station has nothing to transmit, it acts as a simple repeater, that is, it repeats the incoming symbol stream. While repeating, the station must determine if the information is destined for it. This is done by matching the destination address (DA) to its own address or a relevant group address. If a match occurs, then the subsequent symbols up to FCS are processed by the MAC (for MAC frames) or sent to LLC (for LLC frames).

Each transmitting station is responsible for stripping its own transmitted frames from the ring. This is done by stripping the remainder of each frame whose source address (SA) matches the station's address and replacing it with IDLE symbols. This process leaves remnants of frames, consisting of PA, start delimiter (SD), frame control (FC), DA, and SA fields, followed by IDLE symbols, because the decision to remove the frame is based on matching of the SA field and that cannot occur until after the initial part of the frame is repeated. Remnants are easily distinguishable because they are always followed by IDLE symbols. They are removed from the ring when they encounter a transmitting station.

23.3.4. Ring Scheduling

Transmission of normal data on the ring is controlled by a timed token protocol. This protocol supports two types of transfer: (1) asynchronous—dynamic bandwidth sharing and (2) synchronous—guaranteed bandwidth and response time. The asynchronous transfer is typically used for those applications that are bursty and whose response time is not critical. The bandwidth is instantaneously allocated from the pool of remaining unallocated or unused bandwidth. On the other hand, synchronous transfer is the choice for those applications that are predictable enough to permit a preallocation of bandwidth [via station management (SMT)]. Each station maintains a token-rotation timer (TRT) to control the ring scheduling. A target token rotation time (TTRT) is negotiated during ring initialization via the claim token bidding process. Initially, TRT is set to TTRT. A token arriving before the TRT reaches zero, called an early token, causes the TRT to be reset, and can be used for transmitting both synchronous and asynchronous traffic. On the other hand, a token arriving after TRT reaches zero, called a late token, can be used only for synchronous transmissions. Each station also maintains a counter called the late counter (LC), which records the number of times the token was late and a token holding timer (THT), which dictates how long a station can hold the token. This protocol guarantees an average TRT (or average synchronous response time) not greater than TTRT, and a maximum TRT (or maximum synchronous response time) not greater than twice TTRT. When a station receives an early token, that is, LC = 0, it saves the remaining time of TRT in THT and then resets and enables TRT for countdown, that is,

THT = TRT
TRT = TTRT
Enable TRT countdown
Transmit synchronous frames (if any)
Enable THT countdown
Transmit asynchronous frames as long as THT > 0
If the received token is a late token, i.e., LC > 0, then
LC = O
TRT not changed and continues to run
Transmit only synchronous frames

Each station has a known allocation of synchronous bandwidth, that is, the maximum time that the station may hold the token without THT being enabled. Synchronous bandwidth is expressed as a percentage of TTRT; that is, a station would require a 100% allocation to transmit for a time equal to TTRT before issuing a token. The sum of current allocations of all stations should not exceed the maximum usable synchronous bandwidth of the ring (like successive

expiration of TRT with late counter in the MAC or invalid frames on the ring or a logical/physical break in the ring).

23.3.5. Claim Token Process

Any station detecting a requirement for ring initialization shall initiate a claim token process. In this process one or more stations bid (TTRT) for the right to initialize the ring by continuously transmitting claim frames. Each station also looks for incoming claim frames and compares the received bid with the station's own bid. Any station receiving a lower bid shall enter or reenter the bidding, whereas any station receiving a higher bid shall yield. The claim token process is completed when one station receives its own claim frames after the frames have passed around the ring. At this point all stations around the ring have yielded to this station's claim frame. The winning station proceeds to initialize the ring.

23.3.6. Beacon Process

When a station detects that the claim token process has failed or upon a request from SMT, it initiates the beacon process. This happens when the ring has most probably been physically interrupted or reconfigured. The purpose of the beacon process is to signal to all remaining stations that a significant logical break has occurred and to provide diagnostic or other assistance to the restoration process. In this process, the station begins continuous transmission of beacon frames. If a station receives a beacon frame from its upstream neighbor, it repeats that beacon and stops its own. If it receives its own beacon, it assumes that the logical ring is restored and it stops beaconing and starts the claim token process. If a station receives no beacon and it has been continuously beaconing for at least the time indicated by the stuck beacon timer, then ring management begins the stuck beacon recovery procedure. The recovery procedure starts with a directed beacon and eventually traces the stations around the fault, at which time all stations around the fault remove themselves from the ring, perform a self-test, and can rejoin the ring if they pass those tests.

23.4. STATION MANAGEMENT

Station management provides the control necessary at the station (node) level to manage the processes under way in the various FDDI layers such that a station may work cooperatively as part of an FDDI network. SMT provides services such as connection management, station insertion and removal, station initialization, configuration management, fault isolation and recovery, scheduling policies, and collection of statistics. A variety of internal node configurations are possible. However, a node shall have one, and only one, SMT entity. It may, however, have

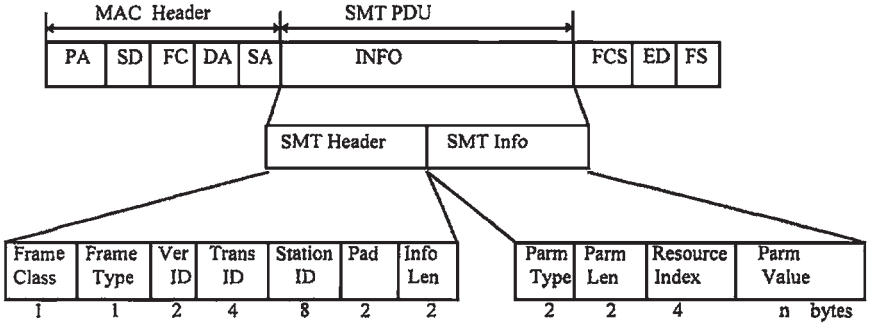


Figure 23.4 Frame subfields.

multiple instances of MAG, PHYs, and PMDs. For a SMT frame, the information field in the general frame format is occupied by a SMT header and a SMT information portion. The SMT header is the protocol header for all SMT frames. SMT information is the information indicated by the SMT header. An SMT frame has subfields as shown in Fig. 23.4.

23.4.1. SMT Header

SMT frames are identified by their frame class, for example, neighbor information frame [NIF(Ol hex)], configuration status information frame [SIFCfg (02 hex)], operation status information frame [SIF-Oper (03 hex)], Echo frame [ECF (04 hex)], and so on. The frame type is an indicator of whether this frame is an announcement (01 hex), a request (02 hex), or a response (03 hex). The ***Version-ID*** field indicates the structure of SMT into field. At the moment, only two possible values are acceptable—0001 hex for stations using a version lower than 7. x and 0002 hex for stations using version 7.x of SMT. To ensure backward compatibility, NIF, SIF, and ECF frames will have a constant ***Version-ID*** of 0001 hex. ***Trunsuction-ID's*** sole purpose is to match the request and response frames. The ***Station-ID*** is the unique identifier of the FDDI station transmitting the SMT frame. The pad field is two bytes of all zeros. This is used to make the header an even 32 bytes. The length of the information field is reported in the info length field. This value does not include the length of the SMT header or this field. Its value can be between 0 and 4458 bytes.

23.4.2. SMT Information

The SMT info field consists of a list of parameters. If more than one parameter is present in the frame, they will be listed one after another. NIFs are used by a station for periodic announcement of its MAC address and basic station

description. NIFs are used in the neighbor notification protocol that allows a MAC to determine its logical upstream neighbor address and logical downstream neighbor address. The protocol also detects duplicated MAC addresses on an operational ring. NIFs can be used by a monitoring station to build a ring map.

SIFs are used, in conjunction with the SIF protocol, to request and provide in response to a station's configuration and operation information. Two classes of SIFs provide this function: the SIF configuration and the SIF operation request and response frames. Potential uses are fault isolation and statistics monitoring.

ECFs are defined for SMT-to-SMT loopback testing on an FDDI ring using the ECHO protocol. The ECH frames may contain any amount of echo data (up to the maximum frame size). Potential uses include confirming that a station's port, MAC, and SMT are operational.

Resource allocation frames are defined to support a variety of network policies such as allocation of synchronous bandwidth. Request-denied frames are used to notify a requester that its request has been denied because of version or protocol problems, for example, if a station receives a request with a ***Version-ID*** that is not supported or if the request frame has a length error.

Status report frames are used by a station to announce the station status, which may be of interest to the manager of the ring. These are sent when certain conditions become (in)active. The conditions that are reported include frame errors, link errors, duplicate address, peer wrap condition, and not-copied condition.

Parameter management frames (PMFs) provide the means for remote access to station attributes via the parameter management protocol. They operate on all SMT MIB attributes. There are two types of PMFs: PMF-Get frames to look up a value and PMF-Set frames to change a value. Not all stations support the PMF-Set protocol.

23.4.3. Physical Connection Management

Within each FDDI station there is one SMT entity per port, called the physical connection management (PCM) entity. The number of PCM entities within a station is exactly equal to the number of ports of that station. PCM entities are a part of SMT that control the operations of the ports. In order to make a connection, two ports must be physically connected to each other by means of fiber-optic or copper cable. When this happens, the PCMs that are responsible for those ports can recognize each other's existence and begin communicating. They do this by sending line states out of the port onto the cable. The PCM at the other end of the connection will recognize the line state and respond accordingly.

When the PCM sees another PCM on the other end of the connection, they will synchronize and communicate with each other. During this communication,

the following important things happen: (i) the PCMs figure out the type of each port and determine if they are compatible and (ii) the PCMs perform a link confidence test (LCT). The LCT determines if the quality of the link is good enough to establish a connection. If not, then a connection will not be made (allowed); otherwise, PCM will establish a connection and place the ports on a token path that goes through that station. At this point the ports become a part of the network and data can be sent through these ports.

23.5. FDDI-II

FDDI-II is an enhancement to the original FDDI being developed by the X3T9.5 committee. Both FDDI and FDDI-II run at 100 Mb/s on the fiber. FDDI can transport both asynchronous and synchronous types of frames. FDDI-II has a new mode of operation called hybrid mode. Hybrid mode uses a 125-ps cycle structure to transport isochronous traffic in addition to synchronous/asynchronous frames. FDDI-II supports integrated voice, video, and data capabilities and therefore expands the range of applications of FDDI. FDDI and FDDI-II stations can be operated in the same ring only in the basic mode.

FDDI-II has the ability to carry the isochronous (constant bit rate) and multimedia data, which is sometimes not possible in the original FDDI due to variation in delay that can be obtained with the token access control method.

FDDI-II nodes can run in basic mode. If all nodes on the ring are FDDI-II nodes, then the ring can switch to the hybrid mode in which isochronous service is provided in addition to basic mode services. In the basic mode on FDDI-II, synchronous and asynchronous traffic is transmitted in a manner identical to that on FDDI. In the basic mode, FDDI-I/I behaves in the same way as FDDI; that is, all ring accesses for transmission are controlled by a token, and the available bandwidth is time shared.

In the alternative mode, also called the hybrid mode, the bandwidth is divided into channels using time-division multiplexing, whereas some of the channels can be reserved for isochronous data. The establishment of multiple channels is done by a station known as the cycle master, which also controls the utilization of the channels. To provide isochronous service, FDDI-II deploys a special frame called a cycle. A cycle is generated every 125 ps. At 100 Mbs, 1262.5 bytes can be transmitted in 125 ps. The bytes of the cycles are preallocated to various channels on the ring. The 1560 bytes of the cycle are divided into 16 wideband channels (WBCs) that can carry either circuit-switched or packet-switched traffic. All packet switched WBCs are concatenated together to form a single channel, which is operated with the FDDI-timed token protocol. Despite the strengths and the characteristics of FDDI-I/I, it has not caught on in the market. Very few vendors have products supporting this standard.

REFERENCES

1. ANSI X3T9.5. FDDI—Physical layer medium dependent (PMD). American National Standards Institute, 1430 Broadway, New York, N.Y. 10018.
2. ANSIX3.148. FDDI-token ring physical layer protocol. American National Standards Institute, 1430 Broadway, New York, N.Y. 10018.
3. ANSI X3.139. FDDI—Token ring media access control (MAC). American National Standards Institute, 1430 Broadway, New York, N.Y. 10018.

24

InfiniBand—The Cluster Interconnect

Ali Ghiasi
Broadcom Corporation

This chapter describes InfiniBand™ (IB) optical link interface. The InfiniBand group was formed to facilitate development of a uniform interconnect from backplane to data center links. Prior to the development of IB, nearly every computer manufacturer developed a proprietary interconnect. Two earlier standardization efforts in this area were SCI (Scalable Coherent Interface) [1] and HiPPi6400 [2], each attempting to develop an open high-performance interconnect. IB takes a further step in defining the high-performance System Area Interconnect by eliminating the bottleneck inside the box. IB link scales from chip, backplane, IO Card, to fiber-optic links. Today InfiniBand is well established in the cluster interconnect, but is not deployed as widely as envisioned due to the market meltdown of 2001, introduction of PCI Express, and 10 Gigabit Ethernet.

24.1. INTRODUCTION

Improvement in VLSI CMOS has enabled fabrication of more complex and faster processors, where the I/O has now become the primary bottleneck [3]. CMOS devices operating at speed greater than 10 Gb/s are widely available [4]. Over the last 10 years wide electrical buses, sometimes 64–256 bit wide operating at a typical speed of 100 MHz, have been replaced with SerDes link operating from 3.125 Gb/s to 10.3125 Gb/s per lane. Further complication with wide slow buses is the associated package size and cost due to high IO count. As the number of I/O increases, additional routing channels are required to route the signals, which increases PCB stack-up layers and the total system cost. It was becoming impractical to increase bus width any longer, and the natural solution has been to increase the CMOS ASIC I/O rate starting with 2.5 Gb/s (SDR), more recently 5.0 Gb/s (DDR) and 10Gb/s (QDR) per I/O.

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

Table 24.1
Performance of InfiniBand Link.

Link Parameters		Raw Bandwidth	
Link Width	Signaling Rate	Unidirectional	Bidirectional
1	2.5 Gb/s (SDR)	250 MByte/s	500 MByte/s
4	2.5 Gb/s (SDR)	1 GByte/s	2 GByte/s
8	2.5 Gb/s (SDR)	2 GByte/s	4 GByte/s
12	2.5 Gb/s (SDR)	3 GByte/s	6 GByte/s
1	5 Gb/s (DDR)	500 MByte/s	1 GByte/s
4	5 Gb/s (DDR)	2 GByte/s	4 GByte/s
8	5 Gb/s (DDR)	4 GByte/s	8 GByte/s
12	5 Gb/s (DDR)	6 GByte/s	12 GByte/s
1	10 Gb/s (QDR)	1 GByte/s	2 MByte/s
4	10 Gb/s (QDR)	4 GByte/s	8 GByte/s
8	10 Gb/s (QDR)	8 GByte/s	16 GByte/s
12	10 Gb/s (QDR)	12 GByte/s	24 GByte/s

Through the link negotiation a low-cost 1X wide 250 MByte/s SDR link can interface to a high throughput 12X wide 12 GByte/s QDR link.

The InfiniBand group early on studied two possible signaling schemes: Source Synchronous and Serial Link. Source Synchronous interfaces have been implemented in SCI with 0.5 Gbyte/s throughput [1], HiPPi6400 at 1 Gbyte/s throughput [2], and PCI Express 8 Gbyte/s throughput [5]. To support backplane and long fiber applications, one has to implement a complex deskew sequence and training similar to HiPPi6400. The alternative based on serial link technology and on 8B/10B, widely in use in Fiber Channel [6] and Gigabit Ethernet [7], provided robust symmetrical signaling, without the need for complex analog de-skewing. To meet high-performance computing (HPC) requirements, a single serial datastream near term from an ASIC could not meet the data throughput. A link layer was developed so that the serial link could be scaled from 1, to 4, 8, and to 12 lanes wide, with each physical lane operating at 2.5 Gb/s (SDR), 5.0 Gb/s (DDR), or 10 Gb/s (QDR). The four-lane implementation has been adapted by IEEE 10 Gigabit Ethernet standard [8] as bases for XAUI interface and similarly by Fiber Channel 10GFC [9].

Infiniband Link provides an interoperable interface with raw bandwidth of 250 MBytes/s to 12 GByte as shown in Table 24.1. Figure 24.1 shows an example of a 4X-SX optical transceiver.

24.2. INFINIBAND LINK LAYER

The Infiniband link layer provides a connection between two InfiniBand protocol aware ports. Each physical link may have 1, 4, 8, or 12 physical lanes. A

Figure 24.1 4X-SX optical transceiver (Courtesy of Avago Technology).

physical link may be copper cable, copper backplane, or fiber optics. Prior to any data transmission, links are trained, width negotiated, and de-skewed.

InfiniBand allows the connection of two protocol aware devices with different width. During link negotiation, the link width common denominator is established between two protocol aware devices, possibly single lane wide. Under an unforeseen circumstance of a lane failing, the link will renegotiate to a new lower common denominator[1] as long as lane 0 is functional. For the case of two 12 lane wide ports, a failure on lane 4 or above results in renegotiation to a 4-lane-wide link.

24.2.1. InfiniBand Packet Format

A protocol aware device sequentially byte strips the data over 1, 4, 8, or 12 lanes, with the start of data packet always on lane 0. Each port that operates from a clock with ±100 PPM allows transmission of 5 Kbytes of uninterrupted data. IB packets including header may be as long as 4608 bytes long, but the actual data portion is only 4096 bytes (4 Kbytes) long.

InfiniBand link uses 8B/10B control characters for packet management [9]. Several key symbols are listed in Table 24.2.

The idle pattern is transmitted on all lanes as the fill pattern to reduce EMI. The pattern is a pseudo-random data generated by an eleventh-order LFSR $X^{11} + X^9 + 1$.

[1]When a link is established, only the following discrete common link widths are supported: 1, 4, 8, and 12.

Table 24.2
Link Control Symbol.

COM	K28.5, Comma, PLL alignment symbol
SDP	K27.7, Start of Data Packet
SLP	K28.2, Start of Link Packet Delimiter
EGP	K29.7, END of Good Packet Delimiter
PAD	K23.7, Packet Padding symbol
SKP	K28.0, Skip Symbol
EBP	K30.7, End of Bad Packet
Idle	Pseudo Random Fill Pattern

The typical packet format for a 4X link is shown in Fig. 24.2. All data transmission starts by start of data packets "SDP" and ends with "EGP" or "EBP." The commas are transmitted for PLL byte alignment. A protocol aware transmitter sends at least three rows of "SKP" to allow clock elasticity adjustment by a maximum of two Retimers in the path. At the input of the protocol aware device, a transmission may have at least one row of SKP or at most five rows of SKP.

The InfiniBand link layer strips the data into a SerDes with 1, 4, 8, or 12 channels for delivery over the fiber or copper media. Figure 24.3 shows the data stripping operation for a 4X link.

24.2.1. IB Electrical Interface

IB electrical interconnect provides typical transmission 20 inches of FR4 printed circuit board (PCB) trace, including two connectors over about 7–17 m of Twin-ax copper cable. The signaling is based on 100 ohms differential impedance with common mode of 0.75 ± 0.45 V for the SDR, and 0.975 ± 0.225 V for the DDR and QDR links. The driver and receiver provide differential source and load terminations of 100 Ω. The minimum differential peak to peak amplitude is 1000 mV for the SDR link, 800 mV for the DDR link, and 600 mV for the QDR link. The receiver sensitivity is 175 mV for the SDR link, 80 mV for the DDR link, and 60 mV for the QDR link. To allow transmission over 20 inches of PCB traces and copper cable, both the preemphasis driver and receive equalizer are required for the DDR and QDR links [11].

IB electrical specifications were designed to be system friendly by relaxing jitter specification to allow ASIC cell implementation. To allow practical ASIC implementation, IB total transmitter jitter is specified at 0.35 (UI) instead of typical 0.25 (UI), which is typically required to drive an optical link. An InfiniBand Retimer or Repeater, sometimes referred to as IB Signal Conditioner [10], is often required prior to direct connection of an IB electrical signal to an

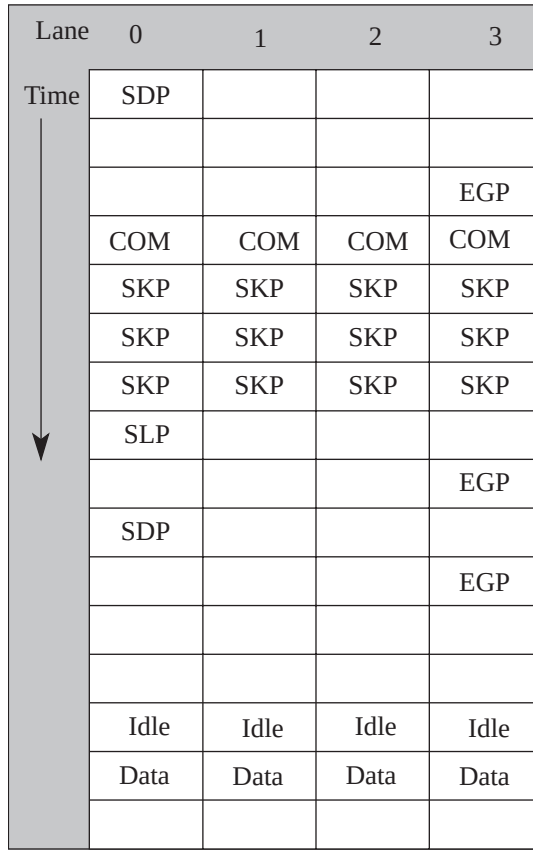


Figure 24.2. Structure of data format as it gets striped across four lanes.

optical transceiver to meet optical jitter specification, due to the relaxed electrical jitter parameters.

A signal conditioner may be implemented with use of CDR retiming or with additional FIFO to allow regeneration from a new reference. The Signal Conditioner may be placed on the board or integrated with the optoelectronic components into a solderable or pluggable assembly for attachment to an IB board. For complete IB electrical specification, see IB High Speed Electrical [11].

24.2.2. Fiber Plant Technology Solutions

IB provides fiber plant options for 1X, 4X, and 12X wide links supporting a broad application base from medium to high performance as listed in Tables 24.3,

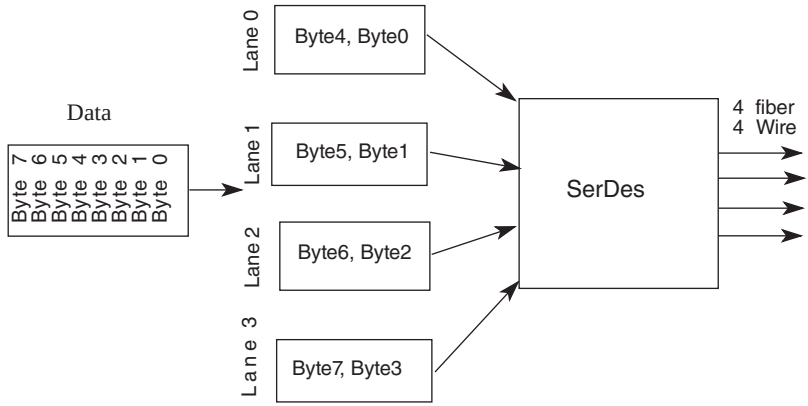


Figure 24.3. Data format striped across 4x link.

24.4, and 24.5, respectively, for the SDR, DDR, and QDR link. Detailed specifications for the three options are provided in Section 24.5.

24.3. OPTICAL SIGNAL AND JITTER METHODOLOGY

This section defines the characteristics of IB compliant optical signals and the measurement methodology. The IB Optical transceiver is composed of an optical transmitter, an optical receiver, and an IB compliant Retimer. The purpose of the IB Retimer is to reduce relaxed IB electrical jitter levels to allow transmission over fiber-optic links. To allow for a scalable architecture, IB electrical signals are designed to drive through 20 inches of FR4 PCB trace with two connectors. At the edge of the PCB, a retimer must regenerate the signal prior to transmission over fiber.

The typical IB compliant link is shown in Fig. 24.4. IB electrical signals are regenerated prior to transmission into the optical transmitter at TP1. TP2 is the optical transmitter output at the optical receptacle. TP3 is the IB complaint point after the designated maximum fiber transmission length as specified in Table 24.3. TP4 is the output of the optical receiver prior to regeneration and transmission as the IB compliant electrical signal levels. It is foreseeable in a self-contained application with tighter electrical specification to interface directly to the optical transmitter and receiver without the use of Retimer. IB optical links are defined such that each lane has a BER $\leq 10^{-12}$ over lifetime, temperature, and operating range when driven through a cable plant which meets IB specifications.

Table 24.3
Fiber-Optic Plant SDR Options.

Link	Very Short Reach (VSR)	Longer Reach
1X Wide		
Designation	IB-1X-SX	IB-1X-LX
Wavelength	850 nm	1300 nm
Connector	dual-LC[1]	dual-LC[1]
Worst-case operating range	2 m–250 m using 50/125 µm 500 MHz.km fiber 2 m–125 m using 62.5/125 µm 200 MHz.km fiber	2 m–10 km with single mode fiber
4X Wide		
Designation	IB-4X-SX	See Note 2
Wavelength	850 nm	
Connector	single MPO	
Worst-case operating range	2 m–125 m using 50/125 µm 500 MHz.km fiber 2 m–75 m using 62.5/125 µm 200 MHz.km fiber	
12X Wide		
Designation	IB-12X-SX	See Note 2
Wavelength	850 nm	
Connector	dual MPO	
Worst-case operating range	2 m–125 m using 50/125 µm 500 MHz.km fiber 2 m–75 m using 62.5/125 µm 200 MHz.km fiber	

1. LC is the registered trademark of Lucent Technology.
2. 4X wide and 12X wide LX may be specified in future versions of the specification.

Table 24.4
Fiber-Optic Plant DDR Options.

Link	Very Short Reach (VSR)	Longer Reach
1X Wide		
Designation	IB-1X-SX	IB-1X-LX
Wavelength	850 nm	1300 nm
Connector	dual-LC[1]	dual-LC[1]
Worst-case operating range	2 m–125 m using 50/125 µm 500 MHz.km fiber 2 m–200 m using 50/125 µm 2000 MHz.km fiber 2 m–65 m using 62.5/125 µm 200 MHz.km fiber	2 m–10 km with single mode fiber

Table 24.4 (*continued*)

Link	Very Short Reach (VSR)	Longer Reach
4X Wide		
Designation	IB-4X-SX	See Note 2
Wavelength	850 nm	
Connector	single MPO	
Worst-case operating range	2 m–75 m using 50/125 µm 500 MHz.km fiber 2 m–200 m using 50/125 µm 2000 MHz.km fiber 2 m–50 m using 62.5/125 µm 200 MHz.km fiber	
12X Wide		
Designation	IB-12X-SX	See Note 2
Wavelength	850 nm	
Connector	dual MPO	
Worst-case operating range	2 m–125 m using 50/125 µm 500 MHz.km fiber 2 m–75 m using 62.5/125 µm 200 MHz.km fiber	

1. LC is the registered trade mark of Lucent Technology.
2. Not currently defined.

Table 24.5
Fiber-Optic Plant QDR Options.

Link	Very Short Reach (VSR)	Longer Reach
1X Wide		
Designation	IB-1X-QDR-SX	IB-1X-QDR-LX
Wavelength	850 nm	1300 nm
Connector	dual-LC[1]	dual-LC [1]
Worst-case operating range	2 m–82 m using 50/125 µm 500 MHz.km fiber 2 m–200 m using 50/125 µm 2000 MHz.km fiber 2 m–33 m using 62.5/125 µm 200 MHz.km fiber	2 m–10 km with single-mode fiber
4X, 8X, 12X Wide		
Designation	IB-4X-QDR-SX See Note 2	IB-4X-QDR-LX See Note 2

1. LC is the registered trademark of Lucent Technology.
2. 4X wide and 12X wide LX may be specified in future versions of the specification.

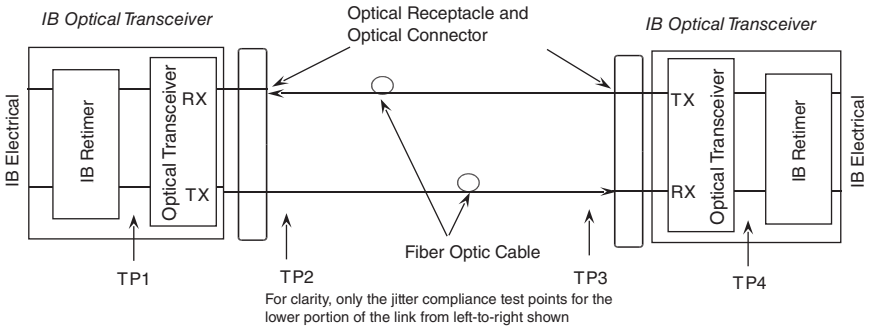


Figure 24.4. IB Optical link and the compliance point.

The Gigabit Ethernet optical link model developed by IEEE 802.3z estimated link performance [6]. However, transmitter power and extinction ratio were replaced by Optical Modulation Amplitudes (OMA) [6] in the link model.

24.3.1. Optical Signal Polarity and Quiescent Condition

An electrical logic zero level at the input of the optical transmitter generates low-level optical power on the fiber, and similarly a logic one generates a high-level optical power.

If electrical input to an optical transmitter is quiescent, then the optical power of that lane is not modulated. IB allows transmitter DC optical power to be left on during a quiescent period.

24.3.2. Optical Transmitter Mask Compliance

The optical transmitter pulse-shape characteristics are specified in the form of a compliance mask on the eye diagram of Fig. 24.5. The eye mask amplitude is normalized so that an amplitude of 0.0 represents logic zero and an amplitude of 1.0 represents logic one. The eye mask normalized timescale is in Unit Interval (UI), where 1U is 400 ps at 2.5 Gb/s. An IB compliant transmitter must meet the eye mask as defined by the reference O/E converter.

This transmitter compliance mask is used to verify the overall response of the optical transmitter for rise time, fall time, pulse overshoot, pulse undershoot, and ringing. Compliance with the optical mask is a good indicator that deterministic effects are within generally acceptable limits, but it does not guarantee compliance with IB jitter specifications.

For uniform eye mask measurements, the optical transmitter signal is measured using an O/E converter with an equivalent fourth-order Bessel-Thompson response given by:

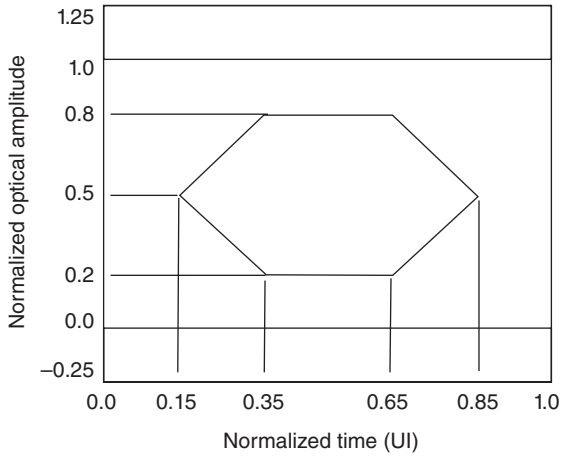


Figure 24.5 Optical transmitter compliance mask.

$$Hp = \frac{105}{105 + 105y + 45y^2 + 10y^3 + y^4}$$

where

$$y = 2.114p \qquad p = \frac{j\omega}{\omega_r} \qquad \omega = 2\pi \cdot f_r \qquad f_r = 1.875\,\text{GHz}$$

The O/E converter filter response is based on the ITU-T G.957 definition, which provides a physical implementation. The specified O/E converter is only intended to provide uniform measurement and does not represent the noise response of an IB Optical Receiver.

The reference O/E converter has a 3-dB frequency response of 1.88 GHz with equivalent response of the fourth-order Bessel-Thompson. The reference O/E converter response must meet the parameters specified in Table 24.6, with tolerance not exceeding the specifications in Table 24.7.

24.3.3 Rise/Fall Times Measurement

Optical rise and fall times are specified based on unfiltered O/E converter waveforms. Some lasers have ringing or overshoot, which can reduce the accuracy of 20%–80% rise and fall time measurements. Therefore, a fourth-order Bessel-Thompson filter defined in a previous section is a convenient filter for measuring rise and fall times. The limited response of the fourth-order Bessel-Thompson filter will adversely impact the measured rise and fall times. The following equation should be used to correct for the filter response:

Table 24.6
Equivalent Response of Reference O/E Converter.

f/f0	f/fr	Attenuation (dB)	Distortion (UI)
0.15	0.2	0.1	0
0.3	0.4	0.4	0
0.45	0.6	1.0	0
0.6	0.8	1.9	0.002
0.75	1.0	3.0	0.008
0.9	1.2	4.5	0.025
1.0	1.33	5.7	0.044
1.05	1.4	6.4	0.055
1.2	1.6	8.5	0.10
1.35	1.8	10.9	0.14
1.5	2.0	13.4	0.19
2.0	2.67	21.5	0.30

Table 24.7
Attenuation Tolerance of Reference O/E Converter.

Reference Frequency f/fr	Attenuation Tolerance Δa (dB)
0.1–1.00	–0.0 . . . +0.5
1.00 . . . 2.00	+0.5 . . . +3.0

Note: Intermediate values of Δa shall be linearly interpolated on a logarithmic frequency scale.

$$T_{rise/fall} = \sqrt{(T_{rise/full}\ \text{measured})^2 + (T_{rise/fall}\ \text{filter})^2}$$

For the purpose of rise and fall time measurement, 3-dB bandwidth of the fourth-order Bessel-Thomson filter may be different from the reference O/E converter with 1.875 GHz bandwidth defined for eye mask compliance.

The IB compliance definition for optical rise time and fall time is RMS of rise and fall times. The IEEE 802.3z Gigabit Ethernet optical link model analysis uses the larger of rise time and fall time, which provides an overly pessimistic analysis [6]. Therefore, IB optical specifications are defined using $T_{r/f}(RMS)$, which is the RMS mean of rise time and fall time as defined in the following:

$$T_{rise/fall\,(RMS)} = \sqrt{\frac{T_{rise}^2 + T_{fall}^2}{2}}$$

24.3.4. Optical Modulation Amplitude

Optical modulation amplitude (OMA) is defined as the absolute difference between the optical power of a logic one level and the optical power of a logic zero level. OMA is related to the extinction ratio (*ER*) measured in (dB) and average optical power (P_{ave} measured in dBm) by the equation:

$$\text{OMA} = 2 \times 10^{\text{Pave}/10} \times \left[\frac{1 - 10^{-\text{Er}/10}}{1 + 10^{-\text{Er}/10}} \right]$$

The OMA specification generally improves the optical transmitter yield as it allows a higher power laser with a lower extinction ratio to pass.

24.3.5. Optical Jitter Specification

The IB jitter specification is based on the same methodology as the Fibre Channel [12] and the IEEE 802.3z Gigabit Ethernet standards [7]. Figure 24.4 shows jitter compliance test points TP1, TP2, TP3, and TP4 for an IB-1X optical link. Similar compliance points exist for the 4X, 8X, and 12X links.

TP1 is an intermediate electrical signal at input of the optical transmitter, with more strength than IB electrical specifications. TP2 is an IB compliant point located at the optical receptacle. Any measurement of optical signal at TP2 is recommended with a 2-meter fiber jumper and wrapped 10 turns around a 25.4-mm diameter mandrel to reduce high-order mode. TP3 is located at the output of the optical fiber over the maximum specified length and adjacent to the optical receiver. TP4 is output of the optical receiver electrical signal. TP1 and TP4 are not IB compliance points and may not be physically accessible.

Typical jitter values at TP1 and TP4 for the SDR and DDR links are listed in Table 24.8 and Table 24.9 for reference. The QDR 1X SX and LX specifications are based on the IEEE 802.3-2002 (Clause 52) standard but with nominal signaling rate of 10.0 Gbaud ±100PPM instead of 10.3125 Gbaud ±100PPM. QDR 1X links are expected to be based on SFP+ [14], which utilizes transmit preemphasis

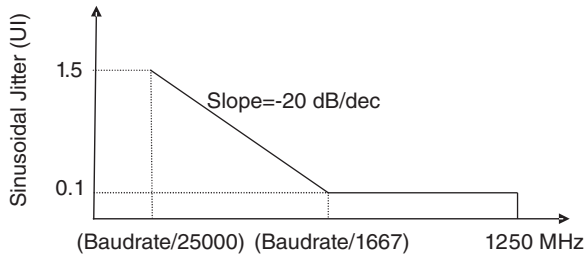


Figure 24.6 Jitter Tolerance Mask.

Table 24.8

SDR Jitter Specifications for Optical Links.

InfiniBand Link	Compliance Point	DJ (UI)	DJ (ps)	TJ (UI)	TJ (ps)
1X-SX,	TP1	0.10	40	0.25	100
1X-LX	TP2	0.23	92	0.46	184
	TP3	0.32	128	0.54	212
	TP4	0.40	160	0.70	280
4X-SX,	TP1	0.10	40	0.25	100
12X-SX	TP2	0.25	100	0.48	192
	TP3	0.30	120	0.53	212
	TP4	0.40	160	0.70	280

Table 24.9

DDR Jitter Specifications for Optical Links.

InfiniBand Link	Compliance Point	DJ (UI)	DJ (ps)	TJ (UI)	TJ (ps)
1X-SX,	TP1	0.14	28	0.26	52
1X-LX	TP2	0.26	52	0.452	90.4
	TP3	0.266	53	0.58	116.1
	TP4	0.365	72.6	0.756	151.1
4X-SX, 8X	TP1	0.10	20	0.25	50
12X-SX	TP2	0.22	44	0.443	89
	TP3	0.22	44.6	0.538	108
	TP4	0.32	64	0.757	151

to compensate for transmit path ISI and a receiver with adaptive equalizer to compensate for receive path ISI. The total jitter of an optical component is measured at BER of 10^{-12} with K28.5+, K28.5− test pattern. An IB optical port must provide BER of $\leq 10^{-12}$ under worst case data patterns and operating conditions. Additional test patterns with detailed methodologies are defined in FC-MJS [13].

The total jitter specified at TP4 in Tables 24.6 and 24.7 does not include any sinusoidal jitter (SJ) component. For all link types, the TP4 must tolerate total jitter of 0.75 UI,[2] with the addition of 0.10 UI of sinusoidal jitter (SJ) over a swept frequency from 1.5 MHz to 1250 MHz.

[2]With additional 0.1 (UI) of SJ the DJ must be reduced by 0.05 (UI) to meet the TJ of 0.75 at TP4.

24.3.6. Jitter Methodology

IB jitter methodology is based on Fibre Channel MJS [13]. Methodologies for Jitter Specification (MJS) breaks down the jitter components into random jitter (RJ) and deterministic jitter (DJ). Random components of successive physical components are added in quadrature, where the DJs are added linearly. Deterministic jitter is composed of duty cycle distortion (DCD), sinusoidal jitter (SJ), and intersymbol interference (ISI). Total jitter is defined as

$$TJ = DJ + RJ = DJ + 14\sigma \qquad \text{for BER of } 10^{-12}$$

The optical jitter compliance points are TP2 and TP3, but values for TP1 and TP4 are provided for reference for the purpose of constructing the SerDes, laser transmitter, and optical receiver. To determine the amount of deterministic jitter and random jitter that an optical transmitter under test adds from TP1 to TP2, the following analysis applies:

$$DJ\,(\text{Transmiter}) = DJ_2 - DJ_1$$

$$RJ\,(\text{Transmiter}) = \sqrt{(TJ_2 - DJ_2)^2 - (TJ_1 - DJ_1)^2}$$

where

DJ1 = DJ at TP1
DJ2 = DJ at TP2
TJ1 = TJ at TP1
TJ2 = TJ at TP2

A similar analysis applies to TP3 and TP4.

24.3.7. Bit-to-Bit Skew

The IB link provides up to 24 ns of skew at the receive input of the protocol aware device between any two receive lanes. To allow interoperable as well as pluggable interfaces, each component has a generous skew allocation. Skew is defined as the maximum differential bit-to-bit skew between any two lanes. All IB optical ports must limit maximum skew to values defined in Table 24.9.

24.4. OPTICAL LINK SYSTEM OVERVIEW

This section deals with the optical link sytem overview for 1X (1 Lane), 4X (4 Lanes), 8X (8 Lanes), and 12X (12 Lanes) wide fiber optic transceivers [15].

Table 24.9
Maximum Optical Bit-to-Bit Skew Values.

Parameter	SDR Skew	DDR Skew	QDR Skew	Note
Optical Cable Assembly	3.0 ns	1.5 ns	0.75 ns	1
Transmitter	500 ps	250 ps	125 ps	2
Receiver	500 ps	250 ps	125 ps	3

1. An optical cable assembly includes optical cable, optical connectors, and adaptors.
2. Between any two physical lanes within a transmitter.
3. Between any two physical lanes within a receiver.

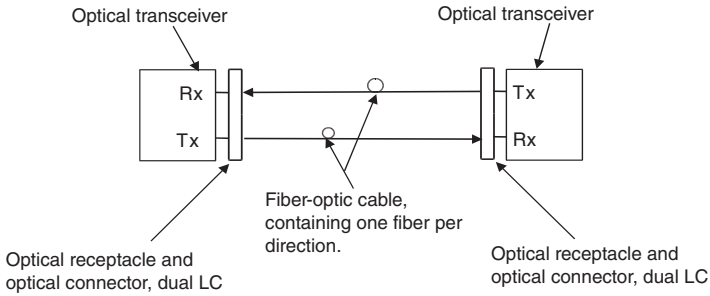


Figure 24.7 1X-SX/LX Fiber-optic link configuration.

InfiniBand optical transceivers may be permanently attached on the circuit board, or they may be fabricated as part of the pluggable through the chassis.

24.4.1. 1X System Overview

The IB 1X optical link carries duplex data as shown in Fig. 24.7; the signals are generated using one laser and one photodetector. Two classes of 1X links are defined: a 1X-SX based on 850-nm VCSEL with multimode fiber for short reach and a 1X-LX based on 1310-nm FP laser with single-mode fiber for operation up to 10 km. IB optical parameters compromised the maximum distance achievable in favor of lower cost and flexibility.

The intermediate reach link, 1X-LX, operates in the 1300-nm wavelength region using single-mode (SM) fiber. The optical parameters are optimized to allow a typical 1X-LX optical transceiver to operate with an uncooled Fabry-Perot (FP) laser, distributed feedback (DFB) laser, or emerging 1310-nm VCSEL [16].

24.4.2. 4X System Overview

A 4X-SX optical link carries duplex data composed of four transmit and four receive signals as shown in Fig. 24.8. The connector is based on a single

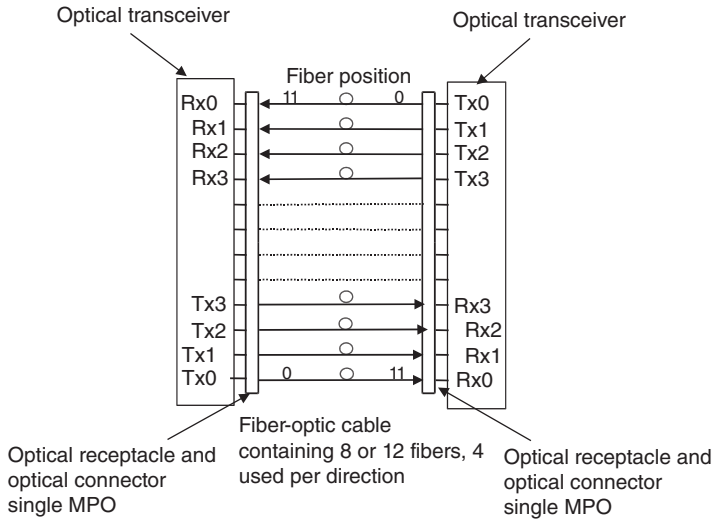


Figure 24.8 4X-SX Fiber-optic link configuration.

12-position MPO[3]; the first four positions carry the transmit signal, followed by four unused positions and the last four positions for the receive signal. To simplify manufacturing and commonality between the 4X and 12X link, ribbon fiber with 12 fibers may be used, but the middle 4 fiber do not carry IB signals. To allow interchangeability between 4X and 12X ribbon cable, the connector keys must face up on both sides of the cable.

24.4.3. 8X-SX System Overview

An 8X-SX optical link carries a duplex 8X data composed of eight transmit and eight receive signals as shown in Figure 24.8, but TX8, TX9, TX10, TX11, RX8, RX9, RX10, and RX11 are absent. The connector is based on a single 12-position dual MPO with four of the fibers unused. The first connector carries eight transmit signals, and the second connector carries eight receive signals. To allow commonality between 4X and 12X, engage the connector key at both ends of a fiber-optic cable with keys up.

24.4.4. 12X-SX System Overview

A 12X-SX optical link carries a duplex 12X data composed of 12 transmit and 12 receive signals as shown in Fig. 24.9. The connector is based on a single

[3]In the United States MPO is known as MTP®, which is a registered trademark of USCONEC Inc.

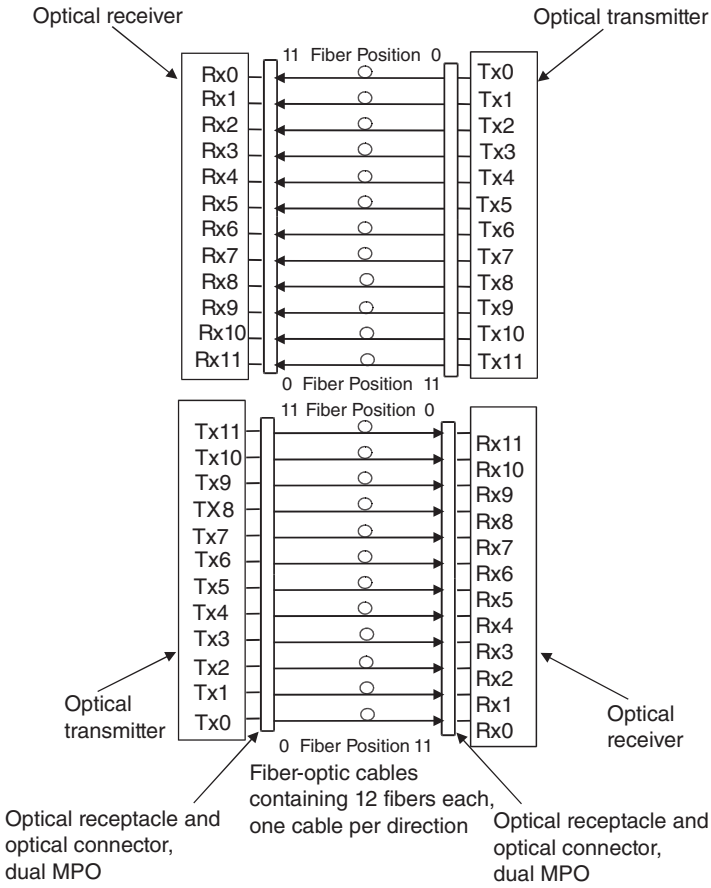


Figure 24.9 12X-SX Fiber-optic Link Configuration.

12-position dual MPO. The first connector carries 12 transmit signals, and the second connector carries 12 receive signals. To allow commonality between 4X and 12X engage the connector key at both end of a fiber-optic cable with keys up.

24.5. OPTICAL RECEPTACLE AND CONNECTOR

This section defines optical receptacle and connector for the 1X link based on LC (LC® Lucent Technologies) and for the 4X/12X link based on the MPO also known as MTP (MTP® USCONEC) in the United States.

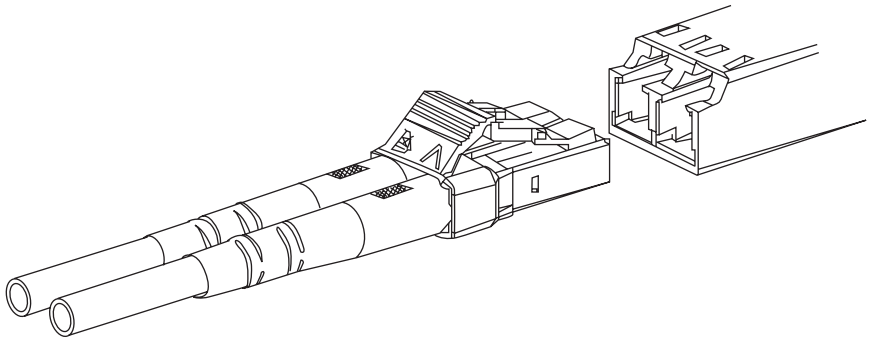


Figure 24.10 Duplex LC plug and socket/receptacle (Courtesy of Systimax Solution).

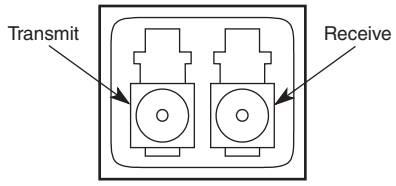


Figure 24.11 1X-SX and 1X-LX optical receptacle orientation looking into the optical port.

24.5.1. 1X Connector—LC

InfiniBand 1X ports are defined based on the Duplex LC optical connector conforming to ANSI/TIA/EIA 604-10 (FOCIS 10), Fiber Optic Connector Intermateability Standard, Type LC. In addition, it will comply with Fibre Channel Physical Interface (FC-PI), Revision 11.0. Figure 24.10 shows a duplex LC connector.

InfiniBand defines orientation for the 1X-SX and 1X-LX fiber-optic transceivers to follow the convention of Fig. 24.11. When looking in to the face plate assuming keys are up, the transmitter is on the left and the receiver is on the right.

To allow ease of distinction between single mode and multimode, the beige color designates multimode and blue designates single mode.

24.5.2. 4X-SX Connector—Single MPO

A single 12-position MPO connector provides bidirectional, full duplex optical connection with single mating action. The 4X-SX optical connector on each end of the fiber-optic cable consists of a female MPO plug conforming to IEC 1754-7-4. Push/Pull MPO Female Plug Connector Interface. The female

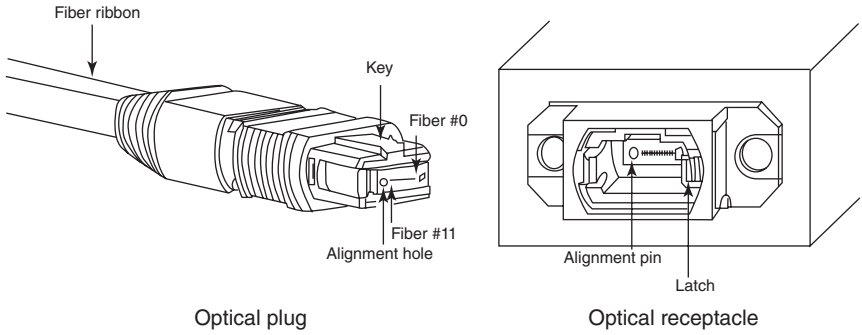


Figure 24.12 MTP optical plug and receptacle (Cortesy of USCONEC).

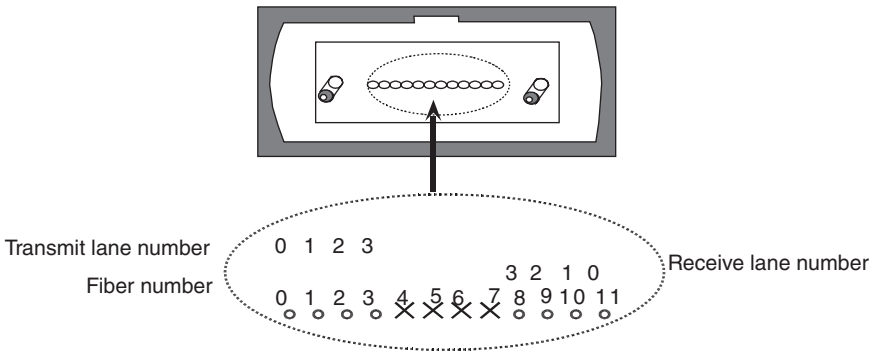


Figure 24.13 Optical receptacle orientation looking into the transceiver.

MPO connector is similar to the male MPO, except that it has no alignment pins. Figure 24.12 shows an outline drawing of an MTP[4] connector plug and receptacle with a push-pull coupling mechanism. Each MPO connector holds a rectangular 6.4 mm by 2.5 mm MT ferrule. The MT ferrule in a male MPO connector holds two precision alignment pins with 0.7-mm diameter.

The 4X-SX optical transceiver interface is a male MPO receptacle. This receptacle has two alignment pins that comply with IEC 1754-7-5 and IEC 1754-7-3; in addition, it conforms to the Push/Pull MPO Adapter Interface standard.

4X-SX optical transceivers follow the transmit/receive convention showed in Fig. 24.13. When looking into the optical receptacle with the key up, fibers are

[4]The IB fiber position definition may not be the same as the manufacture connector fiber position.

numbered from left to right as 0 through 11. The first four fiber positions from the left (fibers 0, 1, 2, 3) are the transmit optical lanes and the last four fibers from the left positions (fibers 8, 9, 10, 11) are the receive lanes. Fibers 0, 1, 2, 3 carry transmit lane 0, 1, 2, 3 signals, respectively. Fibers 8, 9, 10, 11 carry receive lane 3, 2, 1, 0 signals, respectively. The middle four fibers (fibers 4, 5, 6, 7) may physically present but do not carry IB signals.

24.5.3. 8X-SX Connector—Dual MPO

The 8X-SX link is defined based on dual MPO creating a full duplex bidirectional optical connection. The 8X-SX link optical connectors on each end of the cable consist of a two-female MPO plug, conforming to IEC 1754-7-4, Push/Pull Type MPO Female Plug Connector Interface. Each MPO connector plug holds a rectangular 12-position MT ferrule.

The 8X-SX optical port holds two male MPO receptacles, each with two fixed alignment pins conforming to IEC 1754-7-5 and IEC 1754-7-3, Push/Pull MPO Adapter Interface as shown in Fig. 24.12, but transmitter lanes 8, 9, 10, 11 and receiver lanes 8, 9, 10, and 11 are absent. The center-center spacing of the two MPO connectors is 20.0 mm, +/–0.5 mm, allowing development of a duplex plug.

24.5.5. 12X-SX Connector—Dual MPO

The 12X-SX link is defined based on dual MPO creating a full duplex bidirectional optical connection. The 12X-SX link optical connectors on each end of the cable consist of a two-female MPO plug, conforming to IEC 1754-7-4, Push/Pull Type MPO Female Plug Connector Interface. Each MPO connector plug holds a rectangular 12-position MT ferrule.

The 12X-SX optical port holds two male MPO receptacles, each with two fixed alignment pins conforming to IEC 1754-7-5 and IEC 1754-7-3, Push/Pull MPO Adapter Interface as shown in Fig. 24.14. The center-center spacing of the

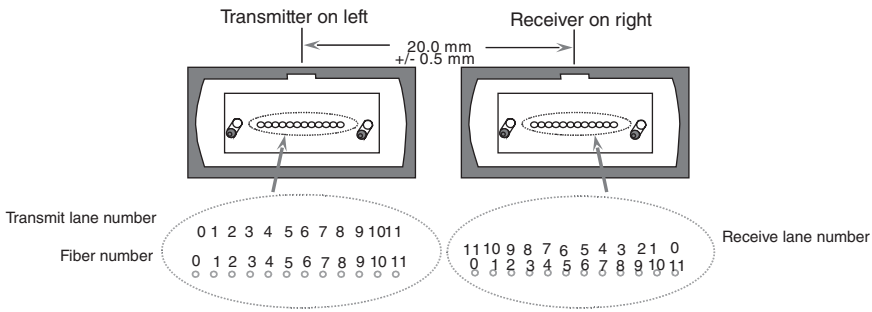

Transmitter on left
Receiver on right
20.0 mm
+/- 0.5 mm
Transmit lane number
0 1 2 3 4 5 6 7 8 9 10 11
Fiber number
0 1 2 3 4 5 6 7 8 9 10 11
11 10 9 8 7 6 5 4 3 2 1 0
0 1 2 3 4 5 6 7 8 9 10 11
Receive lane number

Figure 24.14 12X-SX optical receptacle orientation looking into the optical port.

two MPO connectors is 20.0 mm +/–0.5 mm, allowing development of a duplex plug.

12X-SX optical transceivers follow the transmit/receive convention showed in Fig. 24.14. When looking in to the optical receptacle with key up, the left receptacle carries transmit signals and the right receptacle carries the receive signals. The transmit receptacle fibers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 carry transmit lanes 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, respectively. The receive receptacle fibers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 carry receive lanes 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0, respectively.

24.6. FIBER-OPTIC CABLE PLANT SPECIFICATIONS

This section defines optical fiber cable specifications supporting 1X-SX, 1X-LX, 4X-SX, and 12X-SX variants.

24.6.1. Optical Fiber Specification

IB optical cables comply with the fiber cable specification of Table 24.10 for the respective variant. The single-mode fiber (SMF) conforms to TIA/EIA-492CAAA-98, Dispersion-Unshifted Single-Mode Optical Fibers. Multimode fiber (MMF) 50/125 µm conforms to TIA/EIA-492AAAB-98, Detail Specification for 50-µm Core Diameter/125-µm Cladding Diameter Class Ia Graded-Index Multimode Optical Fibers, or IEC 60793-2 Type A1a. Multimode Fiber (MMF) 62.5/125 µm conforms to TIA/EIA-492AAAA-A-97, Detail Specification for 62.5-µm Core Diameter/125-µm Cladding Diameter Class Ia Graded-Index Multimode Optical Fibers, or IEC 60793-2 Type A1b.

The MMF modal bandwidths specified in Table 24.10 are with overfilled launch condition measured in accordance with the TIA2.2.1 method TIA/EIA/455-204-FOTP204. IB optical system penalties are calculated based on the overfilled launch condition, which gives lower fiber bandwidth. The specification under reasonable conditions is conservative; in practice, better link performance is expected.

Optical passive loss for the cable plant must be verified by the methods of OFSTP-14A. Optical passive loss of a fiber-optic cable is the sum of attenuation losses due to the fiber cable, connectors, and splices.

24.6.2. Fiber-Optic Connectors and Splices

A fiber-optic link may have one or more connectors and splices, provided the total passive loss conforms to the loss budget specified for each varient type. The loss of the fiber plant is verified by the methods of OFSTP-14A. The optical passive loss of a fiber-optic cable is the sum of attenuation losses due to the fiber, connectors, and splices. A total loss budget of 1.5 dB is assigned to optical connectors and splices.

Table 24.10
Optical Fiber Specifications.

Parameters/Fiber Core Supported Variant	SMF (9 µm) 1X-LX	MMF (50/125 µm) 1X-SX,4X-SX, 12X-SX	MMF (62.5/125 µm) 1X-SX,4X-SX, 12X-SX	Units
Conformance	TIA/EIA-492CAAA	TIA/EIA-492AAAB	TIA/EIA-492AAAA	N/A
Nominal Fiber Specification Wavelength	1310	850	850	nm
Fiber Cable Attenuation (Max)	0.5	3.5	3.5	dB/km
Modal Bandwidth with Overfilled Launch (Min)	not applicable	500	200	MHz.km
Zero-Dispersion Wavelength λ_0	$1300 \le \lambda_0 1 \le 1320$	$1295 \le \lambda_0 \le 1300$	$1320 \le \lambda_0 1 \le 1365$	nm
Zero-Dispersion Slope S0 (Max)	0.093	0.11 for $1300 \le \lambda_0 1 \le 1320$ and $0.001\ (\lambda_0 - 1190)$ for $1295 \le \lambda_0 \le 1300$	0.11 for $1320 \le \lambda_0 1 \le 1348$ and $0.001\ (1458 - \lambda_0)$ for $1348 \le \lambda_0 \le 1365$	$ps/nm^2.km$

Fiber-optic connectors and splices for all multimode variants (1X-SX, 4X-SX, 12X-SX) must meet a return loss of 20 dB minimum as measured by the methods of FOTP-107, or equivalent. The single-mode 1X-LX variant must meet a return loss of 26 dB minimum.

REFERENCES

1. Ibel, Maximilian. 1997, August. High-performance cluster computing using SCI. Hot Interconnect V.
2. HiPPi6400, Working Draft, NCITS T11.1, Project 1249-D, Rev 2.2.
3. Charlesworth, Alan. 1997, August. Gigaplane-XB: Extending the ultra enterprise family. Hot Interconnect, Aug 1997.
4. Momtaz, Afshin. Fully integrated SONET OC48 transceiver in Standard CMOS. ISSC 2001, MP 5.2, pp. 76–77. In addition private discussions.
5. PCI Express Card Mechanical Specification, Revision 2.0, April 11, 2007.
6. Fibre Channel Physical Interface (FC-PI). 2001, January. NCITS Project 1235D, Rev 11.
7. IEEE 802.3 Standard. Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications, 802.3z Gigabit Ethernet Section, 2000 Edition (ISO/IEC 8802-3:2000) IEEE Standard for Information Technology.

8. IEEE Draft 802.3ae/D2.3. 2001, March. Media Acess Control (MAC) parameters, physical layer, and management parameters for 10 Gb/s operation. IEEE Standard for Information Technology.
9. Fibre Channel 10 Gigabit (10GFC). 2001, May. NCITS Project 1413-D, Rev 1.1.
10. InfiniBand Architecture Specification, Volume 2, Release 1.2.1, Chapter 5 Link/Phy Interface, October 2006.
11. InfiniBand Architecture Specification, Volume 2, Release 1.2.1, Chapter 6, High Speed Electrical, October 2006.
12. Cunnigham, David. 1999, June. *Gigabit Ethernet networking*. Pearson Higher Education.
13. Fibre Channel—Methodologies for Jitter Specification (MJS), 1999, June. NCITS Project 1230, Rev 10.
14. Next-generation 10 GBaud module based on emerging SFP+ with host-based EDC. Sudeep Bhoja, Ali Ghiasi, Yongmao Frank Chang, Balagopal Mayampurath, Mike Dudek, Shigeru Inano, Eiji Tsumura, 2007, March. *IEEE Communications Magazine* 45, no. 3: 32–38.
15. InfiniBand Architecture Specification, Volume 2, Release 1.2.1, Chapter 8 Fiber Attachment, October 2006.
16. Whitaker, Tim. 2000, July. Long wavelength VCSELs move closer to reality. *Compound Semiconductor* 6(5): 65–67.

This page intentionally left blank

Part IV

Emerging Technologies & Industry Directions

This page intentionally left blank

25

Emerging Technology for Fiber-Optic Data Communication

Chung-Sheng Li
IBM T. J. Watson Research Center

25.1. INTRODUCTION

Due to the explosive growth in the number of Internet users and World Wide Web sites, there has been a significant increase in the demand for the network bandwidth. It has been estimated that there are a total of 1.17 B Internet users (http://www.internetworldstats.com) and 489 M Internet domains (http://www.isc.org) as of 2Q 2007. These Web sites have spawned many academic and commercial applications such as electronic commerce (B2C and B2B), file/music/video sharing (e.g., BitTorrent), Web 2.0 applications (social networks, blogs, wikis, MySpace, YouTube, etc.), VoIP (Vonage, Skype), and IPTV (e.g., Joost). Consequently, the network congestion is aggravated at both the backbone and regional levels in spite of the backbone carriers' effort to increase its bandwidth.

As of August 2007, most of the interconnections between Internet routers at the backbone level operated by UUnet (now a division of Verizon due to the MCI acquisition), British Telecom, AT&T, Sprint Nextel, France Telecom, Quest, Level 3 Communications, and AOL (a division of Time Warner) have mostly been implemented in either OC-48 (2.4 Gbps) or OC-192 (10 Gbps). As the Internet traffic continues its linear growth, the network infrastructure will have serious difficulty catching up with the growth of the traffic, even though there was a temporary excess in capacity as of 2007. Needless to say, future Internet applications such as IPTV (such as Joost), peer-to-peer music/video/file sharing, and Video on Demand (e.g., YouTube) will create even more pressure on the transmission and switching capabilities of the system as more bandwidth-hungry information (graphics, images and video) will be distributed electronically. Therefore, it is natural to expect that the backbone of the future Internet will be based

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

on faster data rates such as OC-192 or beyond and employ some form of optical switching to alleviate the bandwidth problem. As a matter of fact, many carriers have already been in the middle of upgrading these trunk-level capacity to OC-192c (10 Gbps) or even OC-768 (40 Gbps).

Asynchronous Transfer Mode (ATM) over Synchronous Optical Network (SONET) and Multiprotocol Label Switching (MPLS) have already been adopted as the primary transport mechanism for carrying broadband traffic. There is a growing interest in standardizing all the traffic on the IP-based network. In the largest project of its kind, British Telecom announced in 2007 that it will spend 10 billion pounds (about 19 billion U.S. dollars) between 2008 and 2011, replacing its entire public switched telephone network—which serves 22 million subscribers—with a network based on Internet Protocol (IP). The company began its first trial of what it calls 21CN (for twenty-first-century network) with a single exchange in Cardiff, Wales, on November 28, 2006. Roll-out is scheduled to begin in 2008.

Currently, the broadband traffic is carried on single-mode fiber between major switching hubs for data rates up to OC-192 (10 Gb/s). However, the speed of each fiber cannot be increased indefinitely. When the bandwidth required is more than can be supported by a single OC-192 connection, additional multiplexing techniques have to be incorporated in order to advance the link capacity. Currently, as the technology for OC-192 becomes mature, significant research is being devoted to developing OC-768 (40 Gbps) transmission: *Electronic Time-Division Multiplexing (ETDM)* in which four OC-192 channels are time multiplexed together into a single OC-768 channel electronically. In contrast, *Optical Time-Division Multiplexing (OTDM)* multiplexes four OC-192 channels using optical means.

In this chapter, we will survey a number of promising technologies for fiber-optic data communications. The goal is to investigate the potentials and limitations of each technology. The rest of this chapter describes the architecture of all-optical networks including both broadcast-and-select networks and wavelength routed networks; the device aspects of tunable transmitters and tunable receivers for wavelength-division multiplexing (WDM) networks; and the optical amplifier, which is by far the most important technology for increasing the distance between data regeneration so far; wavelength (de)multiplexer technologies; wavelength router technologies; and wavelength converters.

25.2. ARCHITECTURE OF ALL-OPTICAL NETWORK

25.2.1. Broadcast-and-Select Networks

A broadcast-and-select network consists of nodes interconnected to each other via a star coupler, as shown in Fig. 25.1. An optical fiber link, called the carrier, signals from each node to the star. The star combines the signals from all the

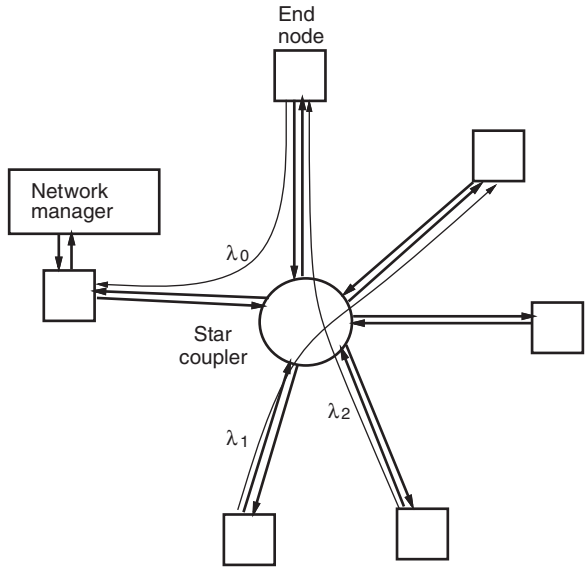


Figure 25.1 A broadcast-and-select network.

nodes and distributes the resulting optical signal equally among all its outputs. Another optical fiber link carries the combined signal from an output of the star to each node. Examples of such networks are Lambdanet [1] and Rainbow [2].

25.2.2. Wavelength-Routed Networks

A wavelength-routed network is shown in Fig. 25.2. The network consists of *static* or *reconfigurable* wavelength routers interconnected by fiber links. Static routers provide a fixed, nonreconfigurable routing pattern. A reconfigurable router, on the other hand, allows the routing pattern to be changed dynamically. These routers provide static or reconfigurable *lightpaths* between end nodes. A lightpath is a connection consisting of a path in the network between the two nodes and a wavelength assigned on the path. End nodes are attached to the wavelength routers. One or more controllers that perform the network management functions are attached to the end node(s).

25.2.3. A Brief WDM/WDMA History

The first field test of the WDM system by British Telecom in Europe dates back to 1991. This test was based on a 5-node, 3-wavelength OC-12 (622 Mb/s) ring around London with a total distance of 89 km. The ESPRIT program, funded by several European governments since 1991, is a consortium funding multiple

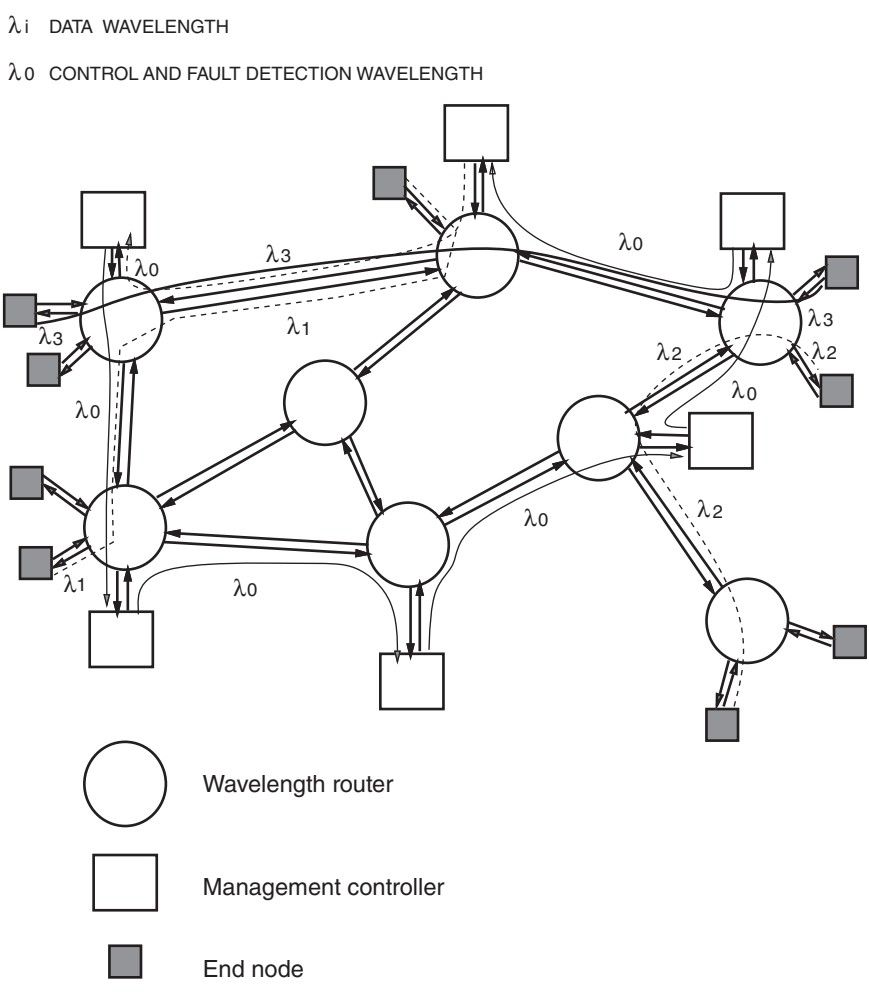


Figure 25.2 A wavelength-routed network.

programs including OLIVES on optical interconnects and several WDM-related efforts. The RACE project, which is a joint university corporate program, has also included demonstrations of the multiwavelength transport network (MWTN). In 1995, ACTS (Advanced Communications Technologies and Services), which includes a total of 13 projects, started building a trans-European Information Infrastructure based on ATM and developing Metropolitan Optical Network (METON). In Japan, NTT Corporation is building a 16-channel photonic transport network with over 320 Gb/s throughput. In the United States, ARPA/DARPA has funded a series of WDM/WDMA activities between 1991 and 1996:

- AON (All Optical Network) consortium, which includes AT&T, DEC, and MIT, focused on developing architectures and technologies that can support point-to-point or point-to-multipoint high-speed circuit-switched multi-gigabits-per-second digital or analog sessions.
- ONTC (Optical Network Technology Consortium), which includes Bellcore and Columbia University, focused on the scalable multiwavelength multihop optical network.
- MONET (Multiple Wavelength Optical Network), which is a consortium including Bell Labs, Bellcore, and three regional Bells, is chartered to develop WDM testbeds and come up with commercial applications for the technology.

Since 1995, many commercially available WDM and DWDM systems have been developed. These systems are described in more detail in Chapter 15.

25.3. TUNABLE TRANSMITTER

The tunable transmitter is used to select the correct wavelength for data transmission. As opposed to a fixed-tuned transmitter, the wavelength of a tunable transmitter can be selected by an externally controlled electrical signal.

Currently, wavelength tuning can be achieved by one of following mechanisms:

- *External cavity tunable lasers*: A typical external cavity tunable laser, as shown in Fig. 25.3, includes a frequency selective component in conjunction with a Fabry-Perot laser diode with one facet coated with antireflection coating. The frequency selective component can be a diffraction grating or any tunable filter whose transmission or reflection characteristics can be controlled externally. This structure usually enjoys wide tuning range but suffers slow tuning time when a mechanical tunable structure is employed. Alternatively, an electro-optic tunable can be employed to provide faster tuning time but narrower tuning range.

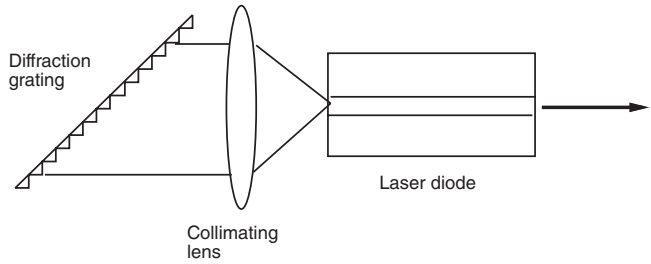


Figure 25.3 Structure of an external cavity laser.

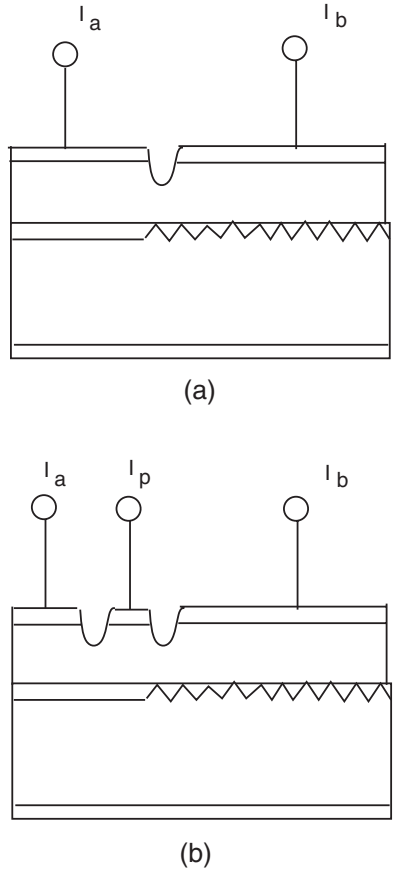


Figure 25.4 Structure of a (a) two-section (b) three-section laser diode.

- *Two-section tunable distributed bragg reflector (DBR) tunable laser diodes*: In a two-section device, as shown in Fig. 25.4, separate electrodes carry separate injection current: one is for the active area, while the other one is for controlling the index seen by the Bragg mirror. This type of device usually has a small continuous tuning range. For example, the tuning range of the device reported in [3] is limited to $\simeq$5.8 nm (720 GHz at 1.55 μm).
- *Three-section DBR tunable laser diodes*: The major drawback of the two-section DBR device is the big gap in the available tuning range. This problem can be solved by adding a *phase-shift* section. With this additional section, the phase of the wave incident on the Bragg mirror section can be varied and matched, thus avoiding the gap of the tuning range.

- *One- or two-dimensional laser diode array [4], or a multichannel grating cavity laser [5], allowing only a few signaling channels.* In another extreme, the wavelengths covering the range of interests can be reached by individual lasers in a one-dimensional or a two-dimensional laser array [4], with each lasing element emitting at a different wavelength. Single dimensional laser array, with each lasing element emitting at single transverse and longitudinal mode, can be fabricated. One possible scheme of single-mode operation can be achieved by means of a short cavity (with high reflectivity coatings) where the neighboring cavity modes from the lasing modes are far from the peak gain. Different emission wavelengths can be achieved by tailoring the cavity length of each array element. For the two-dimensional laser array such as the vertical surface emitting laser array, each array element emits at a different wavelength by tailoring the length of the cavity [4]. It is possible to turn on more than one laser at any given time in both of these approaches. Heat dissipation, however, might limit the number of lasers that can be turned on.

The coupling of the laser emission from the one-dimensional or two-dimensional laser into a single fiber or waveguide can be achieved with gratings or computer generated holographic coupler. Coupling of four wavelengths from a vertical surface emitting laser array has been demonstrated recently. Holographic coupling techniques have also been demonstrated which are capable of coupling over 100 wavelengths into a single link.

Due to the crosstalk and the limited bandwidth of the electronic switch, external modulators might be required to modulate the laser beam for higher bit rate. The light signals can be modulated by either using a directional coupler type modulator, a Mach-Zehnder type modulator [7], or a quantum well modulator [8]. The operation of these devices is required to be wavelength independent over the entire tuning range of the tunable transmitter.

25.4. TUNABLE RECEIVER

The tunable receiver is used to select the correct wavelength for data reception. As opposed to a fixed-tuned receiver, the wavelength of a tunable receiver can be selected by an externally controlled electrical signal.

Ideally, each tunable receiver needs a tuning range that covers the entire transmission bandwidth with high resolution and can be tuned from any channel to any other channel within a short period of time. Tunable receiver structures that have been investigated include:

- Single-Cavity Fabry-Perot Interferometer: The simplest form of a tunable filter is a tunable Fabry-Perot interferometer, which consists of a movable mirror to form a tunable resonant cavity.

  The electric field at the output side of the FP filter in the frequency domain is given by

$$S_{out}(f) = H_{FP}(f)S_{in}(f) \tag{25.1}$$

where H_{FP} is the frequency domain transfer function given by

$$H_{FP}(f) = \frac{T}{1 - Re^{j2\pi\frac{f-f_c}{FSR}}} \tag{25.2}$$

where R is the power transmission coefficient and the power reflection coefficient of the filter, respectively. The parameter f_c is the center frequency of the filter. The parameter FSR is the *free spectral range* at which the transmission peaks are repeated and can be defined as $FSR = c/2\mu L$, where L is the FP cavity length and μ is the refractive index of the medium bounded by the FP cavity mirror. The 3-dB transmission bandwidth FWHM (full width at half maximum) of the FP filter is related to FSR and F by

$$FWHM = \frac{FSR}{F} \tag{25.3}$$

where the reflectivity *finesse* F of the FP filter is defined as

$$F = \frac{\pi\sqrt{R}}{1-R} \tag{25.4}$$

Based on this principle, both fiber Fabry-Perot (as shown in Fig. 25.5) and liquid crystal Fabry-Perot tunable filters have been realized. The tuning time of these devices is usually on the order of milliseconds because of the use of electromechanical devices.

- Cascaded Multiple Fabry-Perot Filters [36]: The resolution of Fabry-Perot filters can be increased by cascading multiple Fabry-Perot filters using either vernier or a coarse-fine principle.
- Cascaded Mach-Zehnder tunable filter [37]: The structure of a Mach-Zehnder interferometer is shown in Fig. 25.6(a). The light is first split by a 3 dB coupler at the input, then goes through two branches with a phase shift

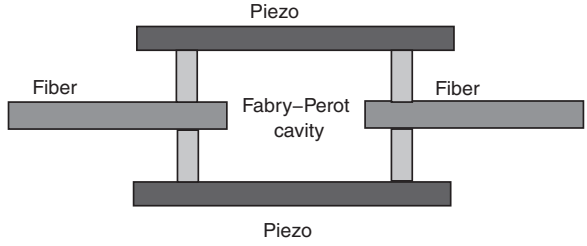


Figure 25.5 Structure of a fiber Fabry-Perot tunable filter.

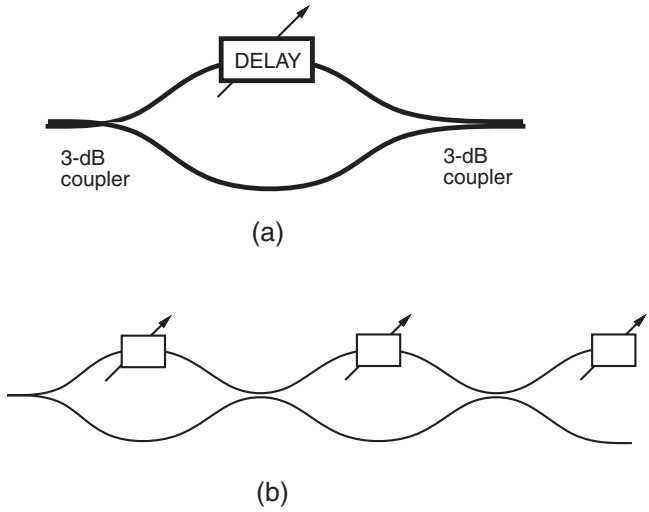


Figure 25.6 (a) Single-stage. (b) Multistage Mach–Zehnder tunable filter.

difference, and then is combined by another 3 dB coupler. The path length difference between two arms causes constructive and destructive interference, depending on the input wavelength, resulting in a wavelength selective device. To tune each Mach-Zehnder, it is only necessary to vary the differential path length by $\lambda/2$. Successive Mach-Zehnder filters can be cascaded together, as shown in Fig. 25.6(b). In order to tune to a specific channel, filters with different periodic ranges will have to be centered at the same location. This can be accomplished by tuning the differential path length of individual filters.

- Acousto-optic tunable filter [37]: The structure of an acousto-optic tunable filter based on the surface-acoustic-wave (SAW) principle is shown in Fig. 25.8. The incoming beam goes through a polarization splitter, separating a horizontally polarized beam from the vertically polarized beam. Both of these beams travel down the waveguide with a grating established by the surface acoustic wave generated by a transducer. The resonant structure established by the grating rotates the polarization of the selected wavelength while leaving the other wavelength unchanged. Another polarization beam splitter at the output collects the signals with rotated polarization to output 1, while the rest is passed to output 2.
- Switchable grating [5]: The monolithic grating spectrometer is a planar waveguide device where the grating and the input/output waveguide channels are integrated as shown in Fig. 25.7. A polarization-independent, 78-channel (channel separation of 1 nm) device has recently been demonstrated

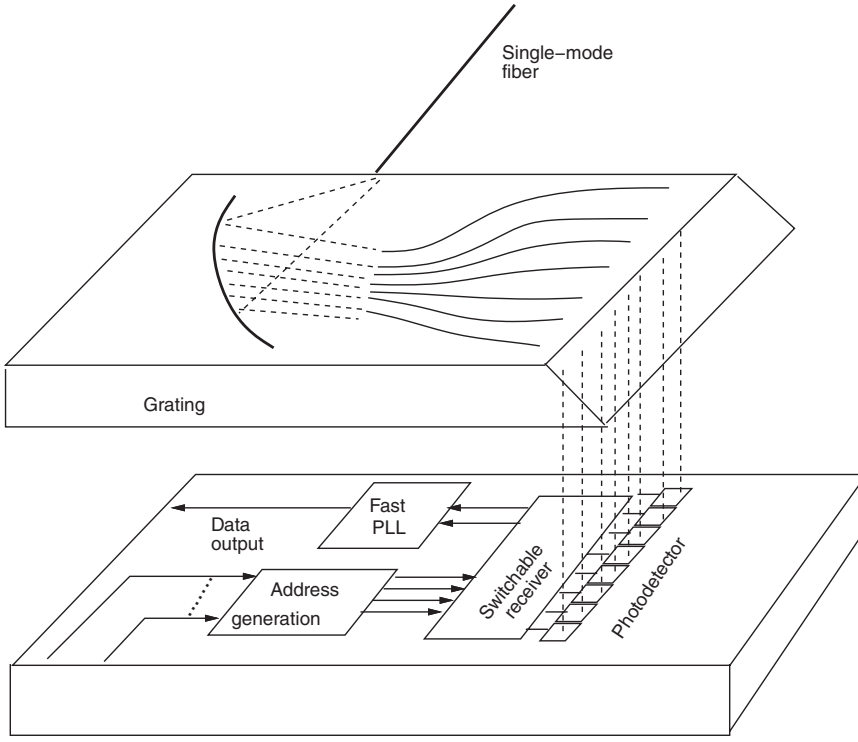


Figure 25.7 Structure of optoelectronic tunable filter using grating demultiplexer and photodetector array.

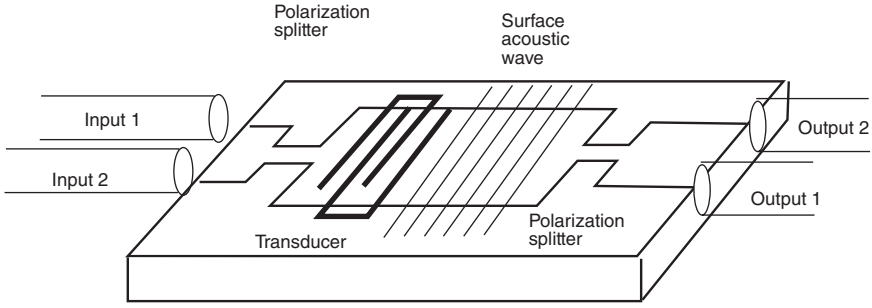


Figure 25.8 Structure of a grating demultiplexer.

with crosstalk ≤–20 dB [5]. For a 256-channel system, the detector array can be grouped into 64 groups with four detector array elements in each bar from which a preamplifier is connected to [9]. The power for the MSM detectors is provided by a 2-bit control as shown. The outputs from the preamplifiers are controlled by gates through a 3-bit control such that every 8 outputs from the gates are fed into a postamplifier. The outputs from the postamplifiers are further controlled by another 3-bit controller. In this way, any channel from the 256 channels can be selected. The switching speed could be very fast, mostly due to the power-up time required by the MSM detectors (the total capacitance need to be driven is ~100 fF ×64).

25.5. OPTICAL AMPLIFIER

In order to achieve all-optical metropolitan/wide area networks (MAN/WAN), optical amplification is required to compensate for various losses such as fiber attenuation, coupling and splitting loss in the star couplers, as well as coupling loss in the wavelength routers. Both rare-earth-ion-doped fiber amplifiers [10, 11] and semiconductor-laser amplifiers can be used to provide amplification of the optical signals.

25.5.1. Semiconductor Optical Amplifier

A semiconductor amplifier is basically a laser diode that operates below the lasing threshold. Two basic types of semiconductor amplifier can be distinguished: the Fabry-Perot amplifier (FPA) and traveling wave amplifiers (TWA). In FPA structures, two cleaved facets act as partial reflective mirrors that form a Fabry-Perot cavity. The natural reflectivity for air-semiconductor facets is 32%, but can be modified through a wide range by using antireflection coating or high-reflection coating. The FPA is less desirable in many applications because of the nonuniform gain across the spectrum. The TWA has the same structure as FPA except that antiflection coating is applied to both facets to minimize internal feedback.

The maximum available signal gain of both the FPA and the TWA is limited by gain saturation. The TWA gain G_s as a function of the input power $P_{in} = P_{out}/G_s$ is given by the following equation:

$$G_s = 1 + \frac{P_{sat}}{P_{in}} \ln \frac{G_o}{G_s} \tag{25.5}$$

where G_o is the maximum amplifier gain, corresponding to the single-pass gain in the absence of input light. It is easy to observe that G_s monotonically decreases to 1 as the input signal power increases, resulting in the gain saturation effect.

Crosstalk occurs when multiple optical signals or channels are amplified simultaneously. Under this circumstance, the signal gain for one channel is affected by the intensity levels of other channels as a result of the gain saturation. This effect depends on the carrier lifetime, which is on the order of 1 ns. Therefore, crosstalk among different channels is most pronounced when the data rate is comparable to the reciprocal of the carrier lifetime.

Another limit to the amplifier gain is due to the amplifier spontaneous emission (ASE). The amplification of spontaneous emission is triggered by the spontaneous recombination of electrons and holes in the amplifier medium. This noise beats with the signal at the photodetector, causing signal-spontaneous beat noise and spontaneous-spontaneous beat noise.

25.5.2. Doped-Fiber Amplifier

When optical fibers are doped with rare-earth ions such as erbium, neodymium, or praseodymium, the loss spectrum of the fiber can be drastically modified. During the absorption process, the photons from the optical pump at wavelength λ_p are absorbed by the outer orbital electrons of the rare-earth ions, and these electrons are raised to higher energy levels. The de-excitation of these high-energy levels to the ground state might occur either radiatively or nonradiatively. If there is an intermediate level, additional de-excitation can be stimulated by the signal photon providing that the bandgap between the intermediate state and the ground state corresponds to the energy of the signal photons. The result would be an amplification of the optical signal at wavelength λ_s.

The main difference between doped-fiber amplifier and semiconductor amplifiers is that the amplifier gain of doped-fiber amplifier is provided by means of optical pumping as opposed to electrical pumping.

Figure 25.9 shows a typical fiber amplifier system. Currently, the most popular doped-fiber amplifiers are based on erbium doping. Similar to semiconductor amplifier, the gain of erbium doped fiber amplifier also saturates. However, the crosstalk effect is much reduced thanks to the long fluorescence lifetime.

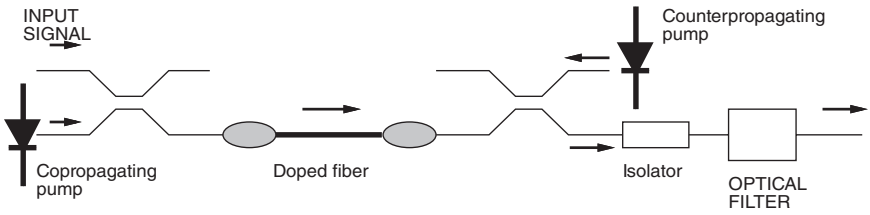


Figure 25.9 A typical doped-fiber amplifier system with either copropagation pump or counterpropagation pump.

25.5.3. Gain Equalization

The gain spectra of these optical amplifiers are nonflat over the fiber transmission windows at 1.3 and 1.55 μm, resulting in nonuniform amplification of the signals. Together with the near-far effect resulting from optical signals originating from various nodes at locations separated by large distances, there exists a wide dynamic range among various signals arriving at the receivers. The best dynamic range of lightwave receivers with high sensitivity reported thus far is limited to less than ~20 dB at 2.4 Gbps [12] and less than ~30 dB at 1 Gbps [13]. In addition, the signal with high average optical power saturates the gain of the optical amplifiers placed along the path of propagation. This limits the available gain for the remaining wavelength channels. Thus, signal power equalization among different wavelength channels is required.

Most of the existing studies on gain equalization have focused on either statically or dynamically equalizing the nonflat gain spectra of the optical amplifiers, but without addressing the near-far effect. For static gain equalization, schemes including grating embedded in the Er^{3+} fiber amplifier [10], cooling the amplifiers to low temperatures [14], or a notch filter [15, 16, 17] were proposed previously to flatten the gain spectra. An algorithm is proposed to adjust the optical signal power at different transmitters to achieve equalization [18]. In [19], gain equalization is achieved by placing a set of attenuators in the arms of the back-to-back grating multiplexers to compensate for nonflat gain spectra of the fiber amplifier. For dynamic gain equalization, a two-stage fiber amplifier with offset gain peaks was proposed in [20] to equalize the optical signal power among different WDM channels by adjusting the pump power. This scheme, however, has a very limited equalized bandwidth of ~2.5 nm. Dynamic gain equalization can also be achieved through controlling the transmission spectra of tunable optical filters. Using this scheme, a three-stage (for 29 WDM channels) [21] and six-stage (for 100 WDM channels) [22] Er^{3+}-doped-fiber amplifier system with equalized gain spectra were demonstrated using a multistage Mach-Zehnder Interferometric filter. An acousto-optic tunable filter has also been used to equalize gain spectra for a very wide transmission window [23]. The combination of these schemes can, in principle, solve the near-far problem in the networks.

25.6. WAVELENGTH MULTIPLEXER/DEMULTIPLEXER

Wavelength multiplexers and demultiplexers are the essential components for constructing wavelength routers. They can also be used for building tunable receivers and transmitters as described in the previous sections.

Two types of wavelength multiplexers/demultiplexers are most widely used: grating demultiplexers and phase arrays.

Figure 25.10 shows an etched-grating demultiplexer with N output waveguides and its cross-sectional view, respectively. The reflective grating uses the Rowland

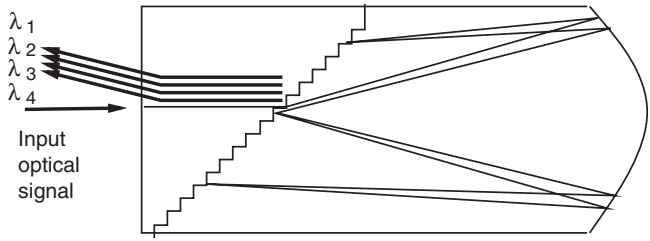


Figure 25.10 Structure of a grating demultiplexer.

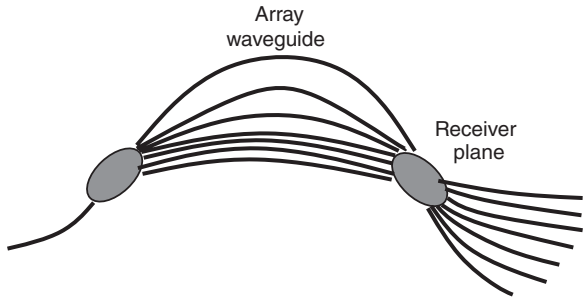


Figure 25.11 Structure of phase array wavelength demultiplexer.

Circle configuration in which the grating lies along a circle, while the focal line lies along a circle of half the diameter.

Phase array wavelength multiplexer/demultiplexers have been shown to be the superior WDM demultiplexers for systems with a small number of channels. A phase array demultiplexer consists of a dispersive waveguide array connected to input and output waveguides through two radiative couplers as shown in Fig. 25.11. Light from an input waveguide diverging in the first star coupler is collected by the array waveguides, which are designed in such a way that the optical path length difference between adjacent waveguides equals an integer multiple of the central design wavelength of the demultiplexer. This results in the phase and intensity distribution of the collected light being reproduced at the start of the second star coupler, causing the light to converge and focus on the receiver plane. Due to the path length difference, the reproduced phase front will tilt with varying wavelength, thus sweeping the focal spot across different output waveguides.

25.7. WAVELENGTH ROUTER

Wavelength routing for all-optical networks using WDMA has received increasing attention recently [24, 25, 26, 27]. In a wavelength-routing network,

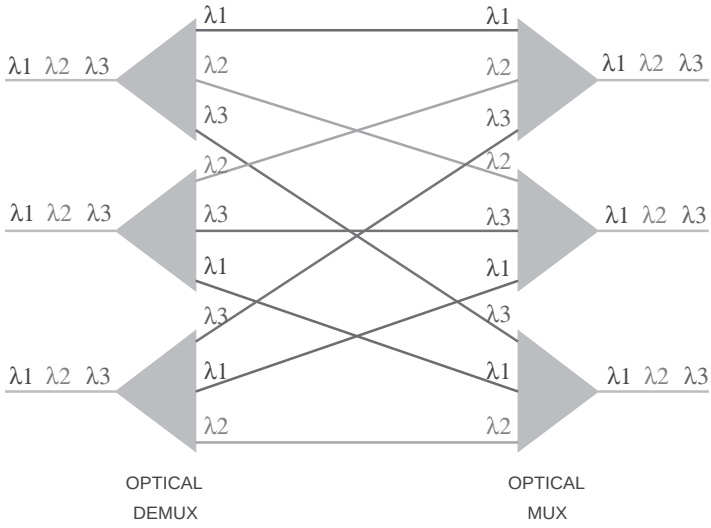


Figure 25.12 Structure of a static wavelength router.

wavelength-selective elements are used to route different wavelengths to their corresponding destinations. Compared to a network using only star couplers, a network with wavelength routing capability can avoid the splitting loss incurred by the broadcasting nature of a star coupler [28]. Furthermore, the same wavelength can be used simultaneously on different links of the same network and reduce the total number of required wavelengths [24].

The routing mechanism in a wavelength router can either be static, in which the wavelengths are routed using a fixed configuration [29], or dynamic, in which the wavelength paths can be reconfigured [30]. The common feature of these multiport devices is that different wavelengths from each individual input port are spatially resolved and permuted before they are recombined with wavelengths from other input ports. These wavelength routers, however, have imperfections and nonideal filtering characteristics that give rise to signal distortion and crosstalk.

Figure 25.12 shows the structure of a static wavelength router, which consists of K optical demultiplexers and multiplexers. Each input fiber to an optical demultiplexer is assumed to contain up to M different wavelengths where $M \leq K$. However, we only consider the case where $M \leq K$. The optical demultiplexer spatially separates the incoming wavelengths into M paths. Each of these paths is then combined at an optical multiplexer with the outputs from the other $M - 1$ optical demultiplexers.

The wavelength routing configuration in Fig. 25.12 is fixed permanently. The optical data at wavelength λ_j entering the *ith* demultiplexer exit at the

[$(j - i)$*modM*]*th* output of that demultiplexer. That output is connected to the *ith* input of the [$(j - i)$*modM*]*th* multiplexer.

Because of the imperfections and nonideal filtering characteristics of the optical multiplexers and demultiplexers, crosstalk occurs in the wavelength routers. On the demultiplexer side, each output contains both the signals from the desired wavelength and that from the other $M - 1$ crosstalk wavelengths. From reciprocity, both the desired wavelength and the crosstalk signals exit at the output on the multiplexer side. Thus, each wavelength at every multiplexer contains $M - 1$ crosstalk signals originating from all demultiplexers.

Crosstalk phenomena in wavelength routers have previously been studied [31, 32, 33, 34]. It was shown in [31] that the maximum allowable crosstalk in each grating (grating as optical demultiplexers and multiplexers in the wavelength router) is −15 dB in an all-optical network with moderate size (say 20 wavelengths and 10 routers in cascade). The results are based on using a 1-dB power penalty criterion and only considering the power addition effect of the crosstalk. Crosstalk can also arise from beating between the data signal and the leakage signal (from imperfect filtering) at the same output channel. The beating of these uncorrelated signals converts the phase noise of the laser sources into the amplitude noise and corrupts the received signals [35] when the linewidths of the laser sources are smaller than the electrical bandwidth of the receiver. Coherent beating, in which the data signal beats with itself, can occur as a result of the beatings among the signals from multiple paths or loops caused by the leakage in the wavelength routers in the system. It was shown in [33] that the component crosstalk has to be less than −20, −30, and −40 dB in order to achieve satisfactory performance for a system consisting of a single, ten, and hundred leakage sources, respectively.

Figure 25.13 shows the structure of the a dynamic wavelength-routing device. This device consists of a total of N optical demultiplexers and N optical multiplexers. Each of the input fiber to an optical demultiplexer contains M different wavelengths. The optical demultiplexer spatially separates the incoming wavelengths into M paths. Each of these paths passes through a photonic switch before they are combined with the outputs from the other $M - 1$ optical switches. When the crosstalk of the optical multiplexer/demultiplexer/switch is considered, each wavelength channel at each input optical demux can reach any of the output optical mux via M different paths.

25.8. WAVELENGTH CONVERTER

The network capacity of WDM networks is determined by the number of independent lightpaths. One way to increase the number of nodes that can be supported by network is to use wavelength routers to enable spatial reuse of the wavelengths, as described in the previous section. The second method is to

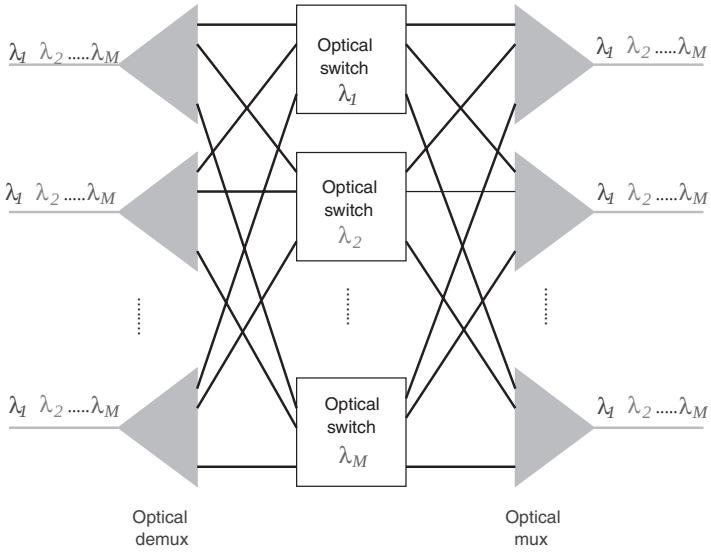


Figure 25.13 Structure of a dynamic wavelength router.

convert signals from one wavelength to another. Wavelength conversion also allows distributing the network control and management into smaller subnetworks and permits flexible wavelength assignments within the subnetworks.

There are three basic mechanisms for wavelength conversion:

1. *Optoelectronic Conversion*: The most straightforward mechanism for wavelength conversion is to convert each individual wavelength to electronical signals, and then retransmit by lasers at the appropriate wavelength. A nonblocking crosspoint switch can be embedded within the O/E and E/O conversion such that any wavelength can be converted to any other wavelengths (as shown in Fig. 25.14(a)). Alternatively, a tunable laser can be used instead of a fixed tuned laser to achieve the same wavelength conversion capability (as shown in Fig. 25.15(b)). This mechanism only requires mature technology. The protocol transparency is completely lost if full data regeneration (which includes retiming, reshaping, and reclocking) is performed within the wavelength converter. On the other hand, limited transparency can be achieved by incorporating only analog amplification in the conversion process. In this case, other information associated with the signals including phase, frequency, and analog amplitude is still lost.
2. *Optical Gating Wavelength Conversion*: This type of wavelength converter, as shown in Fig. 25.15, accepts an input signal at wavelength λ_1 which contains the information and a continuous wave (CW) probe signal at

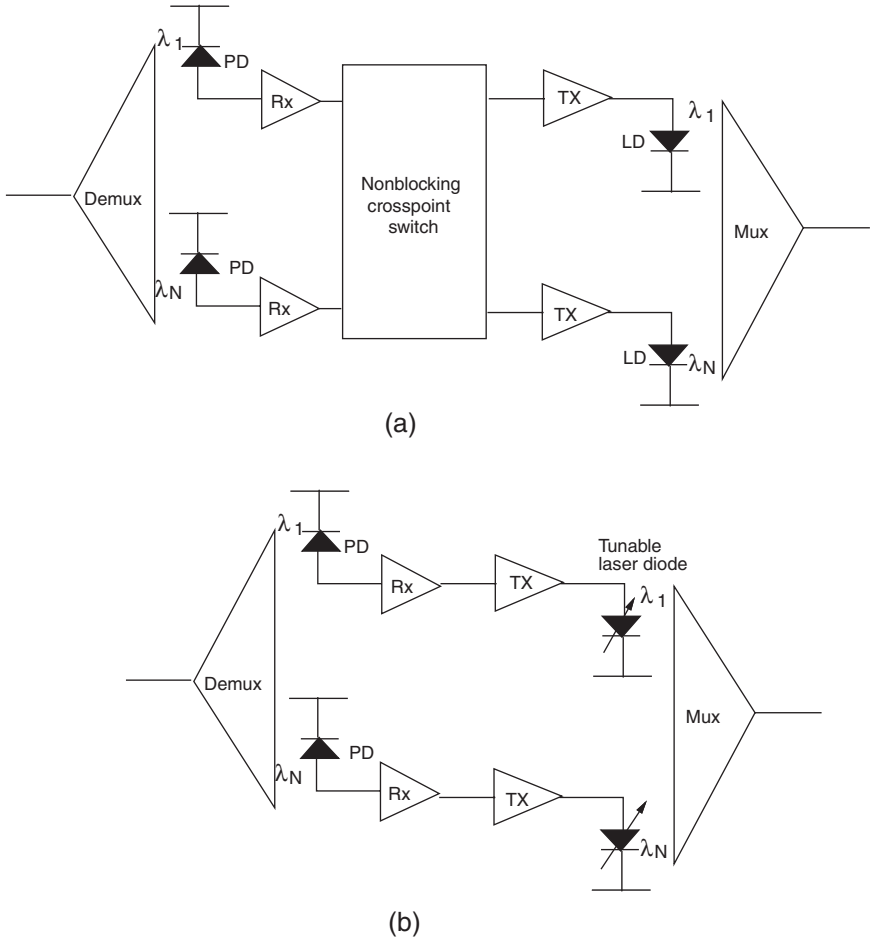


Figure 25.14 (a) Structure of an optoelectronic wavelength converter using electronic crosspoint switch. (b) Structure of an optoelectronic wavelength converter using tunable laser.

wavelength λ_2. The probe signal, which is at the target wavelength, is then modulated by the input signal through one of the following mechanisms:

- Saturable absorber: In this mechanism, the input signal saturates the absorption and allows the probe beam to transmit. Due to carrier recombinations, the bandwidth is usually limited to less than 1 GHz.
- Cross-gain modulation: The gain of a semiconductor optical amplifier saturates as the optical level increases. Therefore, it is possible to modulate the amplifier gain with an input signal, and encode the gain modulation on a separate cw probe signal.

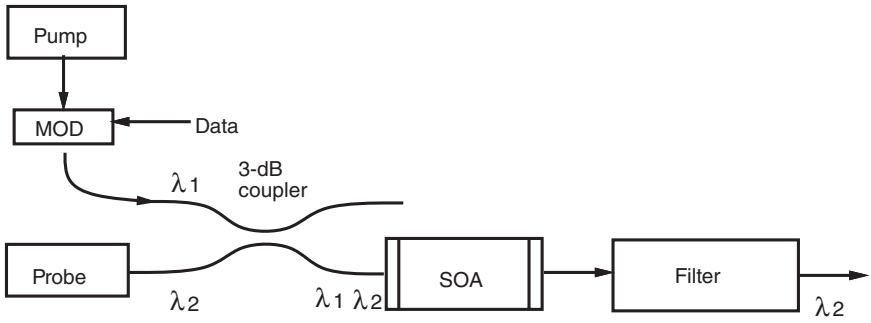


Figure 25.15 Structure of an optical gating wavelength converter using semiconductor optical amplifier.

- Cross-phase modulation: Optical signals traveling through semiconductor optical amplifiers undergo a relatively large-phase modulation compared to the gain modulation. The cross-phase modulation effect is utilized in an interferometer configuration such as in a Mach-Zehnder interferometer. The interferometric nature of the device converts this phase modulation to an amplitude modulation in the probe signal. The interferometer can operate in two different modes: a noninverting mode where an increase in input signal power causes a decrease in probe power, and an inverting mode where an increase in input signal power causes a decrease in probe power.

3. *Wave-Mixing Wavelength Conversion*: Wavelength mixing, such as three-wave and four-wave mixing, arises from nonlinear optical response when more than one wave is present. The phase and frequency of the generated waves are linear combinations of the interacting waves. This is the only method that preserves both phase and frequency information of the input signals and is thus completely transparent. It is also the only method that can simultaneously convert multiple frequencies. Furthermore, it has the potential to accommodate signals at extremely high bit rates. Wave-mixing mechanisms can occur in either passive waveguides or semiconductor optical amplifiers.

25.9. SUMMARY

In this chapter, we have surveyed a number of promising technologies for fiber-optic data communication systems. In particular, we have focused on technologies that can support gigabit all-optical WDM networks. Although most of these technologies are still far from being mature, they nevertheless hold the promise of dramatically improving the network capacity of existing fiber-optical networks.

REFERENCES

1. Goodman, M. S., H. Kobrinski, M. Vecchi, R. M. Bulley, and J. M. Gimlett. 1990, August. The LAMBDANET multiwavelength network: architecture, applications and demonstrations. *IEEE J. Selected Areas in Commun.* 8, no. 6:995–1004.
2. Janniello, F. J., R. Ramaswami, and D. G. Steinberg. 1993, May/June. A prototype circuit-switched multi-wavelength optical metropolitan-area network. *IEEE/OSA J. Lightwave Tech.* 11:777–782.
3. Murata, S., I. Mito, and K. Kobayashi. 1987, April. Over 720 GHz (5.8 nm) Frequency Tuning by a 1.5 μm DBR Laser with Phase and Bragg Wavelength Control Regions. *Electronic Letters* 23, no. 8:403–405.
4. Maeda, M. W., C. J. Chang-Hasnain, J. S. Patel, H. A. Johnson, J. A. Walker, and Chinlon Lin. 1991. Two dimensional multiwavelength surface emitting laser array in a four-channel wavelength-division-multiplexed system experiment. *OFC 91 Digest*: 73.
5. Kirkby, P. A. 1990. Multichannel grating demultiplexer receivers for high density wavelength systems. *IEEE Journal of Lightwave Technology* 8:204–211.
6. Nyairo, K. O., C. J. Armistead, and P. A. Kirkby. 1991. Crosstalk compensated WDM signal generation using a multichannel grating cavity laser. *ECOC Digest*: 689.
7. Alferness, R. C. 1982, August. Waveguide electrooptic modulators. *IEEE Transactions on Microwave Theory and Techniques* 30, no. 8:1121–1137.
8. Miller, D. A. B., D. S. Chemla, T. C. Damen, T. H. Wood, C. A. Burrus, Jr., A. C. Gossard, and W. Wiegmann. 1985, September. The quantum well self-electrooptic effect device: Optoelectronic bistability and oscillation and self-linearized modulation. *IEEE Journal of Quantum Electronics* 21, no. 9:1462–1476.
9. Chang, G. K., W. P. Hong, R. Bhat, C. K. Nguyen, J. L. Gimlett, C. Lin, and J. R. Hayes. 1991. Novel electronically switched multichannel receiver for wavelength division multiplexed systems. *Proc. OFC91*:6.
10. Tachibana, M., R. I. Laming, P. R. Morkel, and D. N. Payne. 1990. Gain-shaped erbium-doped fiber amplifier (EDFA) with broad spectral bandwidth. *Topical Meeting on Optical Amplifier Application*, pp.MD1.
11. Desurvire, E., C. R. Giles, J. L. Zyskind, J. R. Simpson, P. C. Becker, and N. A. Olsson. Recent advances in erbium-doped fiber amplifiers at 1.5 μm. *Proc. Optical Fiber Communication Conference* 1990, San Francisco, Calif.
12. Blaser, M., and H. Melchior. 1992, November. High performance monolithically integrated $In_{0.53}Ga_{0.47}As/InP$ PIN/JFET optical receiver front end with adaptive feed-back control. *IEEE Photonics Technology Letter* 4, no. 11.
13. Mikamura, Y., H. Oyabu, S. Inano, E. Tsumura, and T. Suzuki. 1991. GaAs IC Chip Set for Compact Optical Module of Giga bit Rates. *Proceedings IECON'91 1991 International Conference on Industrial Electronics, Control and Instrumentation.*
14. Goldstein, E. L., V. da Silva, L. Eskildsen, M. Andrejco, and Y. Silberberg. 1993, February. Inhomogeneously broadened fiber-amplifier cascade for wavelength-multiplexed systems. *Proc. OFC'93.*
15. Tachibana, M., R. I. Laming, P. R. Morkel, and D. N. Payne. 1991, February. Erbium-doped fiber amplifier with flattened gain spectra. *IEEE Photonic Technology Letters* 3, no. 2:118–120.
16. Wilinson, M., A. Bebbington, S. A. Cassidy, and P. Mckee. 1992. D-fiber filter for erbium gain flattening. *Electronic Letters* 28:131.
17. Willner, A. E., and S.-M. Hwang. 1993, September. Passive equalization of nonuniform EDFA gain by optical filtering for megameter transmission of 20 WDM channels through a cascade of EDFA's. *IEEE Photonics Technology Letters* 5, no. 9:1023–1026.

18. Chraplyvy, A. R., J. A. Nagel, and R. W. Tkach. 1992, August. Equalization in amplified WDM lightwave transmission systems. *IEEE Photonic Technology Letters* 4, no. 8:920–922.
19. Elrefaie, A. F., E. L. Goldstein, S. Zaidi, and N. Jackman. 1993, September. Fiber-amplifier cascades with gain equalization in multiwavelength unidirectional inter-office ring network. *IEEE Photonics Technology Letters* 5, no. 9:1026–1031.
20. Giles, C. R., and D. J. Giovanni. 1990, December. Dynamic gain equalization in two-stage fiber amplifiers. *IEEE Photonic Technology Letters* 2, no. 12:866–868.
21. Inoue, K., T. Kominato, and H. Toba, 1991. Tunable gain equalization using a Mach-Zehnder optical filter in multistage fiber amplifiers. *IEEE Photonics Technology Letters* 3, no. 8:718–720.
22. Toba, H., K. Takemoto, T. Nakanishi, and J. Nakano. 1993, February. A 100-channel optical FDM six-stage in-line amplifier system employing tunable gain equalizer. *IEEE Photonic Technology Letters*: 248–250.
23. Su, S. F., R. Olshansky, G. Joyce, D. A. Smith, and J. E. Baran. 1992. Use of acoustooptic tunable filters as equalizers in WDM lightwave systems. *OFC Proceeding*: 203–204.
24. Brackett, C. A. 1993. The principle of scalability and modularity in multiwavelength optical networks. *Proc. OFC: Access Network*, p. 44.
25. Alexander, S. B., et al. 1993. A precompetitive consortium on wide-band all-optical network. *IEEE Journal of Lightwave Technologies* 11, no. 5/6:714–735.
26. Chlamtac, I., A. Ganz, and G. Karmi. 1992, July. Lightpath communications: An approach to high-bandwidth optical WAN's. *IEEE Transactions on Communications* 40, no. 7:1171–1182.
27. Hill, G. R. 1988. A wavelength routing approach to optical communication networks. *Proc. INFOCOM:* 354–362.
28. Ramaswami, R. 1993, February. Multiwavelength lightwave networks for commputer communication. *IEEE Communications Magazine* 31, no. 2:78–88.
29. Zirngibl, M., C. H. Joyner, and B. Glance. 1994. Digitally tunable channel dropping filter/equalizer based on waveguide grating router and optical amplifier integration. *IEEE Photonic Technology Letters* 6, no. 4:513–515.
30. d'Alessandro, A., D. A. Smith, and J. E. Baran. 1994, March. Multichannel operation of an integrated acousto-optic wavelength routing switch for WDM systems. *IEEE Photonic Technology Letters* 6, no. 3:390–393.
31. Li, C.-S., F. Tong, and C. J. Georgiou. 1993. Crosstalk penalty in an all-optical network using static wavelength routers. *Proc. LEOS Annual Meeting*.
32. Li, C.-S., and F. Tong. 1994. Crosstalk penalty in an all-optical network using dynamic wavelength routers. *Proc. OFC'94*.
33. Goldstein, E. L., L. Eskildsen, and A. F. Elrefaie. 1994, May. Performance implications of component crosstalk in transparent lightwave networks. *IEEE Photonics Technology Letters* 6, no. 5:657–660.
34. Goldstein, E. L., and L. Eskildsen. 1995, January. Scaling Limitations in Transparent Optical Networks Due to Low-level Crosstalk. *IEEE Photonic Technology Letters* 7, no. 1:93–94.
35. Gimlett, J., and N. K. Cheung. 1989. Effects of phase-to-intensity noise conversion by multiple reflections on gigabit-per-second DFB laser transmission systems. *IEEE Journal of Lightwave Technologies* 7, no. 6:888–895.
36. Hamdy, W. M., and P. A. Humblet. 1993. Sensitivity analysis of direct detection optical FDMA networks with OOK modulation. *IEEE Journal of Lightwave Technologies* 11, no. 5/6:783–794.
37. DeCusatis, C., and P. Das. 1991. *Acousto-optic signal processing fundamentals and applications*. Boston: Artech House.

This page intentionally left blank

Case Study Customer-Owned Wavelengths and P2P Optical Networking

In the early history of computing, large centralized data centers were the repository of mainframes and storage devices, typically connected to networks of "dumb" terminals. Computer architectures required the construction of larger and more powerful computers with each successive generation. This centralized model of computing was fundamentally changed with the adoption of client-server architectures. (Indeed, this shift in computing models led many analysts to incorrectly predict the end of mainframe computing.) Large enterprise servers have continued to provide value in sizable customer accounts worldwide. Many technologies have been developed for the control of both datacom and telecom optical networks in a client-server environment, including Generalized Multi-protocol Label Switching (G-MPLS) and the Optical User Network Interface (O-UNI). While these protocols have been adopted in carrier-class telecom networks, the client-server approach to networking is not without its weaknesses. For example, customers cannot independently change the bandwidth or topology of a virtual private network (VPN), which has emerged as a desirable requirement in modern grid computing (where large, bursty traffic exists between widely separated locations). Likewise, since optical VPN services are only edge-to-edge, users on different VPNs cannot be easily cross-connected within a single carrier network or between multiple independent management domains. Under the client-server model, a client domain does not allow other domains to access its internal links, or the links between itself and its service providers. However, the service provider offers all its client domains access to all end users in the network. Thus this approach is highly asymmetric in resource utilization.

By contrast, peer-to-peer (P2P)-based flat network architectures exist widely in the Internet and distributed computing applications. These networks are

characterized by participants who share part of their hardware resources (processing power, storage capacity, etc) in order to enable various service offerings (file sharing or shared workspaces). In such networks, the users may also be viewed as service and content providers, directly accessible by their peers without intermediate organizations. It is also possible to construct various hybrid networks with characteristics of both P2P and client-server architectures (for example, the well-known Napster music sharing service is a P2P model, which also employs a central node as the file directory). Recently, technologies have begun to emerge, which allow client domains to build their transport networks using P2P networks. This has implications for the network physical layer. While all peer domains may be considered equal, they do not necessarily share the same physical connectivity; some domains may be able to perform only fiber cross-connects, and others, individual wavelength switching or packet-based add/drop of SONET frames. There may be different transmission or switching bandwidth requirements at the edge nodes of a P2P network than near its core.

Dark fiber has become available in some metropolitan or wide area networks to facilitate domain sharing. With the widespread deployment of WDM optical networks, however, it has become possible to consider new services such as leased wavelengths. Although leased wavelengths are still mostly engineered by service providers, the end-to-end optical path used for domain interconnection can be established by the domains themselves. This allows functionally different domains to be connected with the same features as dark fiber, but with more flexibility and potentially lower cost. There is ongoing research in this area, including problems related to the improvement of interdomain routing and signaling protocols for sharing control and management data. One interesting example is the national research and education network funded by the government of Canada and supported by the nonprofit development group CANARIE (http://www.canarie.ca/about/index.html). Under development for many years, this network's ultimate goal is to allow end users to provision, manage, and control their own guaranteed bandwidth connections on a P2P basis. Operating without the need for end users to request service from a centralized network management authority, the end users are primarily research institutions or universities pursuing bandwidth-intensive research such as grid computing projects in high-energy physics, bio-informatics, and related fields. While the current network retains some hybrid features, similar to most industry standard telecom systems, the concept of end users controlling the setup, teardown, and routing of their own connections has made some significant progress. Rather than claiming P2P will completely replace existing networks, a better analogy might be the introduction of the private branch exchange (PBX) in telephone systems. Although similar services were available from both the PBX and centrally managed telephone exchange, the PBX allowed users to manage their own internal connectivity (for example, using tie-line phone numbers within their facility). Over time, the PBX

also tended to be the first to introduce new, innovative services and aggressive pricing. Just as the centralized telecom market evolved, it is expected that optical network protocols for large carrier networks will continue to be required in addition to P2P; this may present new opportunities for technologies able to exploit the best features of both approaches.

REFERENCES

Many references on the Canadian national network and peer-to-peer networking can be found at: http://www.canarie.ca/canet4/library/canet4design.html

Halabi, B. 1997. *Internet routing architectures.* Indianapolis, Ind.: Cisco Press, New Riders Pub.

Parameswaran, M., A. Susarla, and A. B. Whinston. 2001. P2P networking: An information-sharing alternative. *IEEE Computer* 34:31–36.

Saleh, A. A. M., and J. M. Simmons. 1999. Architectural principles of optical regional and metropolitan access networks. *Journal of Lightwave Technology* 17:2431–2448.

Schollmeier, R. 2001, August 27–29. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. *Proceedings of the 2001 International Conference on Peer-to-Peer Computing*. Linkopings University, Sweden, pp. 101–102.

Wu, J., J. M., Savoie, and B. St. Arnaud, 2002, June. Peer to peer optical networking—architecture, functional requirements, and applications. World Scientific (http://www.worldscientific.com/), 3CFD01C2-5963-22C4.doc.

This page intentionally left blank

26

Optical Backplanes, Board and Chip Interconnects

Rainer Michalzik
Ulm University, Institute of Optoelectronics, Ulm, Germany

26.1. INTRODUCTION

This chapter is a completely revised version of a corresponding previous book chapter [1]. Related information on, for example, vertical-cavity surface-emitting laser (VCSEL) -based frame-to-frame optical interconnects using 10 Gbit/s-grade high-bandwidth multimode fibers, parallel optical links, or plastic optical fibers has been omitted, and the focus is now entirely on intrasystem interconnects at the backplane, printed circuit board, and chip levels. Since the list of references has been substantially shortened and updated, the readers might find it helpful to consult [1] in some instances. Apart from general remarks on the technologies, we will illustrate approaches to practical interconnect systems by means of examples taken from the recent literature, mainly of the years 2004 to 2007. Because of the highly dynamic nature of this field of research and the limited text length, necessarily no complete overview can be given here.

The main motivation for research into ultra-short-distance optical links bridging distances in the 1 m- to the cm-range is the same as 20 years ago, namely, the obvious problems with electrical lines operated at ever increasing clock frequencies of electronic processing systems. In particular, these are resistance–capacitance (RC) and inductance–capacitance (LC) related time constants, both incurring a limited transmission bandwidth, ohmic losses, electrical crosstalk between adjacent lines, or experienced as electromagnetic interference in general, traveling wave reflections due to impedance mismatch, ground bounce, increasing power consumption, connector handling due to high insertion force at large pin counts, and so on. These problems are contrasted by seemingly obvious solutions provided by optics, namely, absence of transmission line effects, wider bandwidth, reduced crosstalk which is even frequency-independent, much improved

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

data rate scalability of data buses, voltage isolation, higher channel density and thus much higher bandwidth density, connectors with low insertion force, and so on. From this comparison one might conclude that the incorporation of optics within electronic systems should progress at a rapid pace. However, the current state-of-the-art is greatly lagging behind previous predictions, according to which optical backplanes should already be a commercial reality [2]. Thus the purpose of the present text is not to only praise the advantages of optical solutions but also to clearly point out their weaknesses.

The following section focuses on optical backplanes, which, according to current implementations, are purely passive in the sense that they do not incorporate optical signal generation, reception, or processing capability. Instead, they serve to distribute optical signals between numerous boards within a rack, and thus both platforms require proper connectorization. Whereas on an optical backplane, waveguiding is mostly envisioned as taking place within an optical layer located at the top surface, boards realistically require that the optical layer be buried within the multiple electrical layers. Board-level optical interconnects are thus treated in a separate section. There is a transition region between board interconnects and optical interchip communication taking place within a single board without any intended connection to the backplane. Such interconnects as well as the extreme case of optical intrachip links are contained in the fourth section of this chapter.

26.2. OPTICAL BACKPLANES

26.2.1. General Remarks

It is very reasonable to assume that within the next five to ten years, optical signals will gradually penetrate into individual boxes of high-end systems like telecom switches, Internet servers, or computer clusters. This will happen at the level of optical backplanes in combination with some intraboard transmission. In [1], approaches to optical backplane interconnects are subdivided into the categories (1) discrete cabling, (2) integrated waveguides, (3) free-space, straight path, (4) free-space, zigzag path, and (5) direct interboard. These are separately discussed, and numerous references are given. The first category employs optical fibers or fiber ribbons routed between transceiver modules arranged at the board edges, which requires special types of connectors. The components do exist, and the approach is used in practical systems; however, it hardly satisfies the desires for compactness, high-bandwidth density, and low cost. Categories three to five use free-space optics with stringent requirements on alignment. Since hardly any progress has been achieved over the last years, these approaches can be ruled out to find practical applications in the foreseeable future. Thus, in this section we focus entirely on the integrated waveguide type of optical backplane, which has reached a high degree of maturity.

An overview of optical backplane transmission containing lots of useful references is given in [3]. The emphasis of the paper is on the power budget, analyzing that the inclusion of signal routing schemes necessitates the use of local optical amplification. For this purpose, a novel active device is proposed which can potentially be flip-chip mounted on the backplane. Since at present virtually all demonstrators in the literature exclusively employ point-to-point links, optical amplification is not further considered in this text. Concerning the waveguide attenuation coefficient, an upper limit of 0.05 dB/cm is deemed acceptable, such that the propagation loss remains below 5 dB for link lengths up to 1 m [4]. Whereas this value is easily reached with optical fibers, for polymer waveguides it excludes using the established telecommunication wavelengths of 1.31 and 1.55 µm since with presently favored materials, losses are at least in the 0.3- to 0.5-dB/cm range [3]. Instead, most demonstrators run at the standardized optical datacom wavelength of 850 nm, where high-performance and low-cost VCSELs, as well as photodiodes, are available [1].

In [5] it is pointed out that higher bandwidth is the most compelling motivation for using board-level optical interconnects. With a somewhat idealized analytical model, partition lengths are found beyond which optical interconnects can outperform their electrical counterparts in terms of aggregate bandwidth. These lengths depend on the available semiconductor technology generation, have a tendency to decrease within decreasing feature size, and are predicted to be in the 10-cm range in the year 2010, which is well within the range of board-level transmission. A different modeling approach including critical characteristics of the link-end devices is taken in [6]. The comparison of power consumption yields link lengths of a few 10 cm, which favor optical solutions. There is some uncertainty in the given numbers owing to numerous assumptions that have to be made. Nevertheless, such analyses present a strong motivation for considering the use of optics on the backplane and board levels.

26.2.2. Example Backplane Demonstrators

A particularly practical approach to an optical backplane from [7] was presented in [1] and is briefly repeated here. Figure 26.1 shows its main features, namely, expanded beam free-space coupling at the board-to-backplane interfaces, which eliminates the need for optical connectors, 90° beam deflection by micromirrors, and polymer channel waveguiding within the backplane. Large lateral alignment tolerances of ±0.5 mm for 1-dB excess loss were reached solely by aid of the remaining electrical pins required for power and low-speed control lines. The waveguides have losses of as low as 0.03 dB/cm. For the specific targeted application in avionics, immunity to electromagnetic interference was the main concern, such that the low bandwidth density enforced by the few mm expanded beam diameter was judged acceptable.

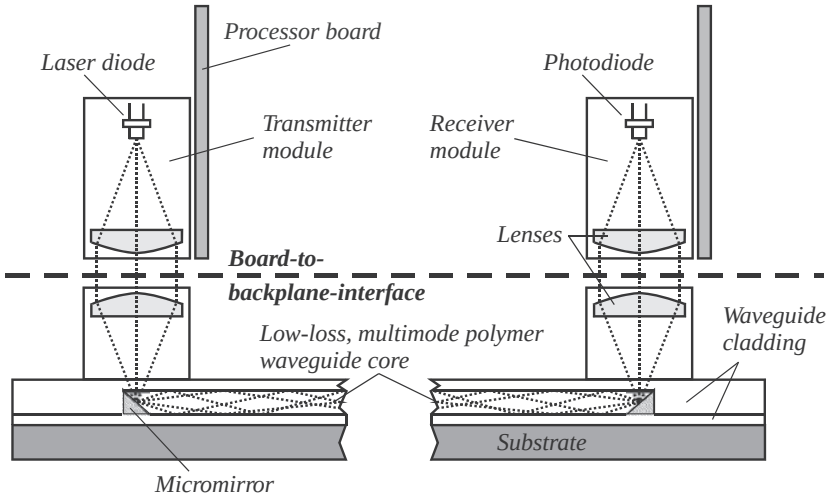


Figure 26.1 Optical backplane concept using free-space coupling [7].

A completely different approach to coupling the printed circuit boards (PCBs) to the optical backplane is illustrated in Fig. 26.2. In this case, special optical plugs were designed and fabricated [8]. They incorporate a 45° mirror and a microlens to couple the signal into the waveguides, which are laminated into the backplane board. The backplane is slotted, and the boards are inserted and aligned by means of precision optical adaptors. Signal propagation predominantly occurs in polymer waveguides with a loss of 0.1 dB/cm at 850 nm wavelength, which were created with hot embossing. However, multimode fiber sections are also used in the optical plug. In [1], several previous projects employing in-board fibers are referenced; a more recent example is given in Section 26.3.2. Although glass fibers in particular have negligible loss over board-relevant distances, their assembly is labor-intensive, signal path bending is difficult, there is no cost-reducing planar fabrication and packaging technology, and fiber termination is rather challenging.

Figures 26.1 and 26.2 nicely illustrate a main shortcoming of current optical backplane systems, namely, an optical board-to-backplane interface that is compatible with existing rack technology. To be able to compete with electrical solutions, desired features of a connector should include easy insertion, high-density, two-dimensional waveguide arrangements, and good manufacturability leading to low cost. Although numerous concepts are found in the literature, present inefficiencies seem to be a major hurdle for the gradual introduction of optical interconnects at the board and backplane levels.

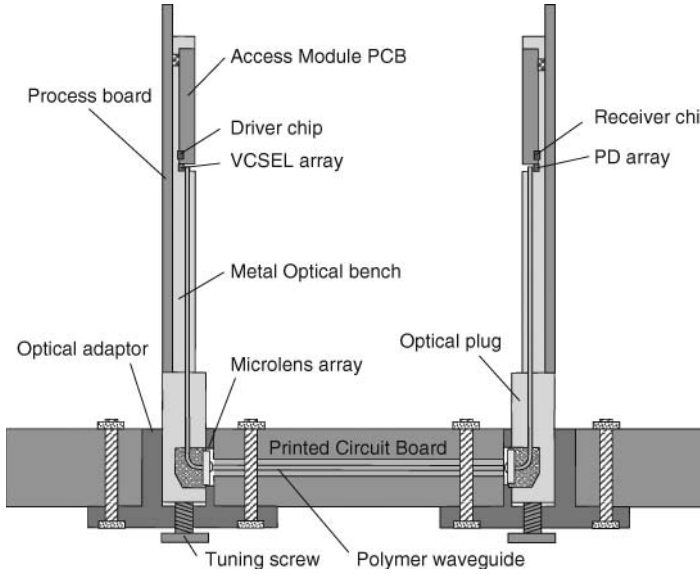


Figure 26.2 Schematic of an optical backplane system using connection slots [8] (© 2004 IEEE).

26.2.3. Integrated Polymer Waveguides

As mentioned before, free-space optical approaches to backplane- and board-level interconnects are impractical and excluded from the present discussion [1]. Instead, except for a few experiments using embedded fibers, polymer channel waveguides are the preferred photon transport medium. Ideally, one or several optical layers containing the waveguides are included in the PCB, which additionally has a multitude of electrical layers. On the other hand, as an intermediate step, several chip-to-chip interconnect demonstrators nowadays still use waveguide arrays deposited on the board surface (see Section 26.4.2). In this section we will summarize some important topics concerning such polymer waveguides.

First, it is clear that multimode rather than single-mode waveguides are preferred, since they offer relaxed alignment tolerances but still provide sufficient bandwidth for less than 1 m transmission length. The waveguides usually have a quadratic cross section with side lengths in the 30- to 100-μm range. Waveguide pitches in linear arrays are mostly 250 μm—compatible with multimode fiber ribbons—but can be made much smaller owing to negligible optical crosstalk. Several polymers with attenuation coefficients in the 0.1-dB/cm range at 850-nm wavelength are reported in [9]; see also [3]. Eligible material classes are acrylates,

siloxanes, polyimides, polycarbonates, and olefins. Often, the exact materials choice is kept confidential in optical interconnect publications.

To find acceptance in the established electronics industry, waveguide integration in an embedded optical layer should be fully compatible with today's PCB manufacturing sequence. In particular, the following requirements have to be met. The waveguide layer has to withstand high pressure (applied during lamination, for example, 1.5 to 4 MPa over 1 h for FR-4 material) at elevated temperature (e.g., 160°C) without deterioration especially of the optical properties. High temperatures also occur during the soldering step. Lead-free soldering in an infrared lamp-heated oven will increase the temperature up to 260°C for a period of up to 20 s, which accordingly indicates the desired glass transition temperature of the polymers. Preferably the optical layer should be placed in the center of the PCB to avoid undesired warp and at the same time reduce the temperature load. As stated before, for longer distance links the attenuation coefficient of the waveguides should not exceed 0.05 dB/cm. For waveguide fabrication, several options exist. (1) Photolithography with an ultraviolet (UV) lamp provides good resolution, can produce low-loss waveguides, is suitable for mass production, and can nowadays even handle usual board sizes. (2) UV direct write (e.g., using a He:Cd laser) also gives very low losses, large sample sizes are possible, and there is large flexibility; however, as a serial technique (with some room for parallelization), it cannot yield high throughput. (3) Hot embossing employs a master tool, is well suited for mass production, and can directly integrate 45° surfaces used as mirrors but incurs a higher surface roughness (leading to loss coefficients exceeding 0.1 dB/cm) and, due to the high pressure required, can handle only limited board sizes. From this comparison it is concluded that photolithography would be the method of choice for industrial production, which of course requires photopatternable polymer materials.

As an example, Fig. 26.3 shows an acrylate-based waveguide array. High-density arrays with 100 μm pitch and $35 \times 35\,\mu m^2$ core size were fabricated [4] and loss coefficients in the 0.035 to 0.05 dB/cm range were achieved at 850 nm wavelength. Crossings of 30-μm-wide waveguides have losses as low as 0.02 dB. For out-of-plane light coupling, 45° mirrors are often used. Fabrication

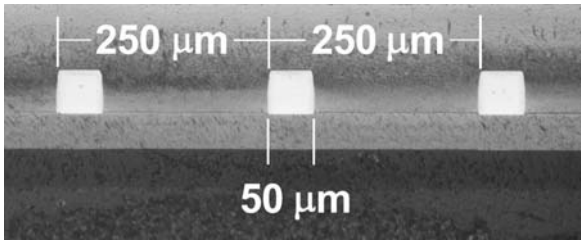


Figure 26.3 Linear polymer waveguide array with 250-μm pitch and $50 \times 50\,\mu m^2$ cross-sectional area (Courtesy of IBM Research).

technologies include microdicing with a V-shaped diamond blade [10], laser ablation [11], reactive ion etching [12], and hot embossing (see before), among others (see also [3] and the references therein). In general, since minimum bending radii for tolerable attenuation usually are in the millimeter range and the incorporation of low-loss 90° deflections often proves difficult, planar optical waveguide circuits are rather space consuming. More options for optical signal routing are reported in [13].

26.3. OPTICAL BOARD INTERCONNECTS

26.3.1. Coupling to Intraboard Waveguides

For the interfacing of transmitter and receiver elements with the waveguides, essentially three options are discussed, as illustrated in Fig. 26.4. In the first case (a), transceiver modules (i.e., at least the optoelectronic device arrays and the corresponding electronic driver and receiver chips) are placed on top of the board, where light from the VCSEL is guided down to the waveguide and back up to the photodetector (PD) at the receiving end. Deflections by 90° can be implemented either in the guiding pin or the waveguide itself. Alternatively, free-space propagation is possible. In the second option (b), laser and PD are displaced from

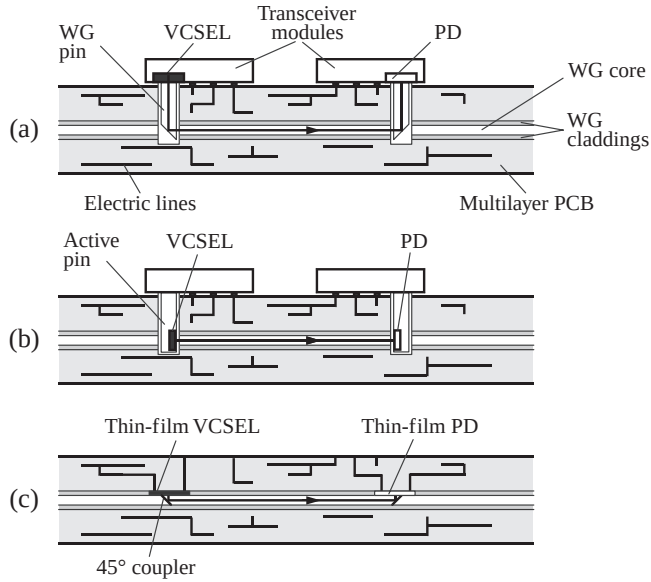


Figure 26.4 Options for interfacing optoelectronic devices to the waveguide (WG) layer in a hybrid electrical–optical printed circuit board (PCB). Indirect coupling via waveguide pins (a), direct butt-coupling (b), and board-embedded one-dimensional VCSEL and photodetector (PD) arrays (c).

the transceivers such that they butt-couple with the waveguide. The third case (c) envisions thin-film optoelectronic devices fully embedded into the optical layer, where electrical contacting is provided through the conventional multilayer structure. All three approaches have found practical implementations. Option (b) is rather critical in the sense that the active devices can easily be damaged during module insertion and is thus not much followed. Some work in this direction is referenced in [1]. For illustration, few recent results relying on options (a) and (c) are presented in the next section.

26.3.2. Example Board Interconnect Demonstrators

Figure 26.5 shows a coupling approach equivalent to Fig. 26.4(a) [14]. The optical connection rods are segments of 250 µm pitch, 12-channel silica multimode fiber ribbons with 62.5 and 100 µm diameter at the transmitter and receiver ends, respectively. Mirror planes at an angle of 45° were formed by mechanical polishing. Rod insertion into the via holes with about 140 µm diameter, formed by CO_2 laser drilling, was facilitated by guide pins in a plastic ferrule holding the fibers. The embedded polymer waveguide cores are 100 µm wide and 65 µm high. The total loss from VCSEL to photodiode amounts to 8 dB. A data rate of 2.5 Gbit/s could be transported over 5 cm interconnection length. In the actual experiment, separate transmitter and receiver boards were used and the VCSELs and photodiodes were placed at a distance of about 10 µm from the flat surface of the rods, where an arrangement as in Fig. 26.5 is, however, feasible. The rods are envisioned to be mass producible by plastic molding, and their insertion should be possible with modern pick-and-place equipment. Through different lengths of the connection rods, the scheme should be applicable to two-dimensional interconnects over stacked waveguide layers.

One of the few demonstrations of actual two-dimensional optical board interconnection is reported in [15]. As schematically illustrated in Fig. 26.6, a 7.5-cm-long 2×6 array of silica optical fibers with 100-µm core diameter and the usual

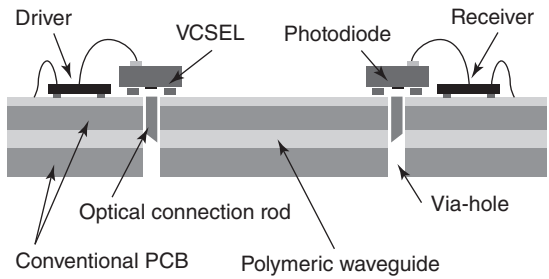


Figure 26.5 Schematic of a board-level optical link using optical connection rods for light coupling [14] (© 2004 IEEE).

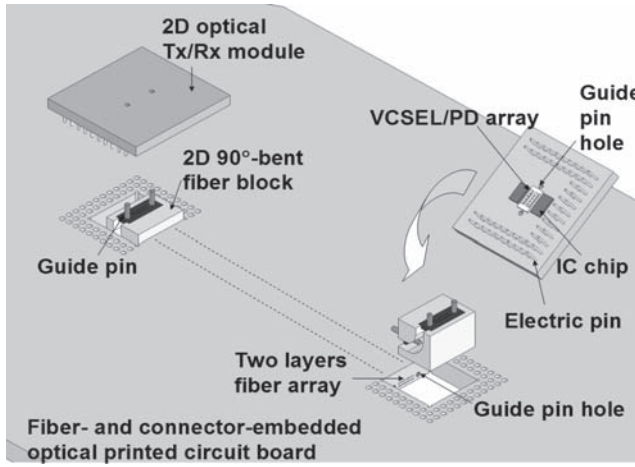


Figure 26.6 Layout of a two-dimensional board-level interconnect scheme using embedded fibers and bent-fiber connection modules [15] (© 2007 IEEE).

250-μm pitch is embedded in a four electrical layer PCB, where the optical layers have a distance of 500 μm and the fibers terminate in modified MT ferrules. The signals are directed to the surface via $6 \times 8 \times 6\,\text{mm}^3$ large MT-pluggable modules incorporating two fiber layers, which are bent by 90° with radii of curvature of 3 and 3.5 mm. VCSEL and photodiode arrays as well as driver and receiver chips are mounted such that each stacked fiber pair acts as transmit and receive channel. The components and their passive assembly with the aid of guide pins are described in [15] in some detail. The optical insertion loss is 5.3 dB, with index matching oil at the board-to-connector interfaces, and 3 Gbit/s data transmission is shown.

Fiber embedding in boards is attractive from the propagation loss perspective and is investigated in [16] in some detail. Fibers are placed in grooves produced with a dicing saw. A large board prototype is successfully fabricated under realistic conditions. Nevertheless, fiber endface quality after via hole formation, and thus coupling loss and interchannel uniformity, remain major concerns. Moreover, only rather slight bends might be incorporated.

Entirely embedding the optoelectronic devices into the PCB as in Fig. 26.4(c) is a particularly attractive vision. Ideally, the board performance is enhanced by buried optical interconnects without any extra space consumption. Figure 26.7 shows conceptual drawings illustrating the work in an ongoing project [12]. Current demonstrations incorporate 1×12 arrays of thin-film VCSELs and pin-type photodiodes as well as 45° total internal reflection micromirrors. Laser reliability is one of the major concerns, and thermal device performance and optimization

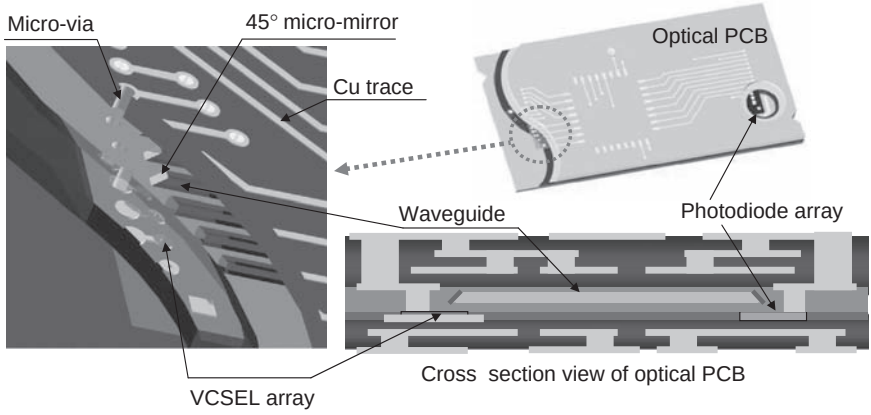


Figure 26.7 Schematic of a fully embedded board-level optical interconnect system [17] (© 2006 IEEE).

are studied in [17, 18]. PCB packaging compatibility and sufficient lifetime still need to be proven. Thin-film VCSELs are also presented in [19]; however, both optical waveguides and lasers are to be placed at the board surface, similar to some chip-to-chip interconnect examples in Section 26.4.2. Waveguide-embedded optoelectronic devices within postprocessed optical layers on a PCB are discussed in [20]. In this scheme, edge-emitting laser diodes and lateral coupling to metal–semiconductor–metal photodetectors is favored, thus entirely avoiding beam turning elements.

26.4. OPTICAL CHIP INTERCONNECTION

In this section we treat both interchip and intrachip optical links. While the interchip links have close relationships with (usually) waveguide-based board-level interconnects and appear rather practical, intrachip links are mostly implemented using free-space optics and are much more speculative.

26.4.1. General Remarks

Many publications on chip-level interconnects refer to the Semiconductor Industry Association's roadmaps [21], which predict a severe bottleneck for off-chip and global on-chip interconnections to be encountered in future technology generations. Insufficiencies in electrical interconnects especially lead to high power consumption and obstructed processor–memory interaction. The numerous physical reasons to use photons instead of electrons for interconnection on principle also apply at the chip scale (see [22] and the references therein). However, there

exist many roadblocks that prevent a mainstream use of optical chip interconnects, among which are chip real estate requirements for optoelectronics-related circuits, skepticism to any postprocessing of fabricated VLSI chips, high operating temperatures in the 100°C range that require extraordinary reliability of active optoelectronic devices, lack of computer-aided design tools merging the electrical and optical worlds, and finally the cost issue [1]. Moreover, unforeseen advances are continously being made in electrical chip interconnection (such as proximity communication [23]) which make the penetration of optics into the chip level even more unlikely. With or without optics, bandwidth and latency are the two fields desiring the most progress. While in the bandwidth case, dense two-dimensional arrangements of optoelectronic devices hold some promise for huge amounts of interchip data throughput, possible reductions of signal latency are much less obvious when considering signal propagation through additional driver and receiver circuits (see, e.g., [24] for architectural implications).

26.4.2. Example Interchip Interconnects

Figure 26.8 shows an interconnect approach based on so-called active interposers [25]. These are ceramic platforms that carry driver or receiver integrated circuits (ICs) at their upper surface and linear arrays of 850-nm top-emitting VCSELs or $100 \times 100\,\mu m^2$ active area metal–semiconductor–metal photodiodes (PDs) at their bottom surface, where connection is made with electric through-holes. In the experiment, the module size is $12 \times 11 \times 0.67\,mm^3$, and there is ample space for, among other things, logic circuits (denoted as LSI in the figure). With a target clearance of $60\,\mu m$ between the optoelectronic chips and the waveguide layer placed on top of the board, light coupling to a 10-channel waveguide array with $250\,\mu m$ pitch and $40 \times 40\,\mu m^2$ core size is possible without any optics via uncoated 45° total internal reflection mirrors. Assembly on the PCB is done through passive alignment. Data rates of 2 Gbit/s have been successfully transmitted between interposers. Sealing of the modules as well as coupling to buried waveguides still needs to be addressed.

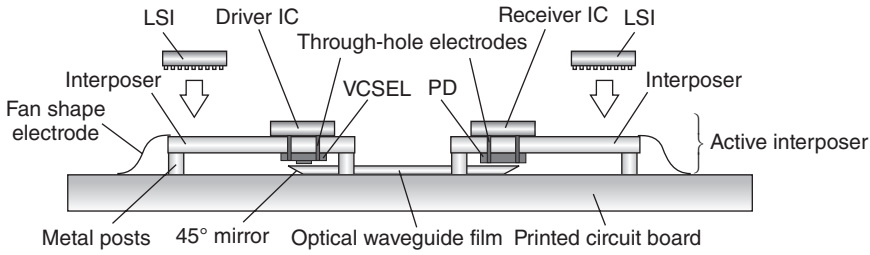


Figure 26.8 Schematic cross-sectional view of a chip-to-chip link based on active interposers at the transmitting and receiving ends [25] (© 2006 IEEE).

Alternative board-level interconnect packages are presented in [26], which are compatible with standard surface-mount technology (SMT). To account for misalignments, microlens arrays are incorporated both at the bottom of the packages and at the top of the PCB, which image the optoelectronic device arrays onto the 50-μm core size buried waveguides, as schematically illustrated in Fig. 26.9. In the demonstrator, three channels are arranged at 500-μm pitch, and the 480-μm diameter lenses form 400-μm diameter collimated beams, such that even misalignments up to ±100 μm of the ball grid array package produce losses of not more than 1 dB. Data transmission experiments at 1.25 Gbit/s show wide-openeye diagrams. The approach has some similarity to the board-to-backplane coupling scheme in Fig. 26.1 and might be limited by the achievable bandwidth density.

Significant advances in board-level module-to-module optical interconnection have been made in the U.S. Terabus project [27], which has set state-of-the-art targets for individual channel data rates of 20 Gbit/s and for both high aggregate throughput and bandwidth density using 48 parallel polymer waveguides placed at 62.5-μm pitch. Major applications are foreseen in future server environments. An overview of the developed package is given in Fig. 26.10. The so-called

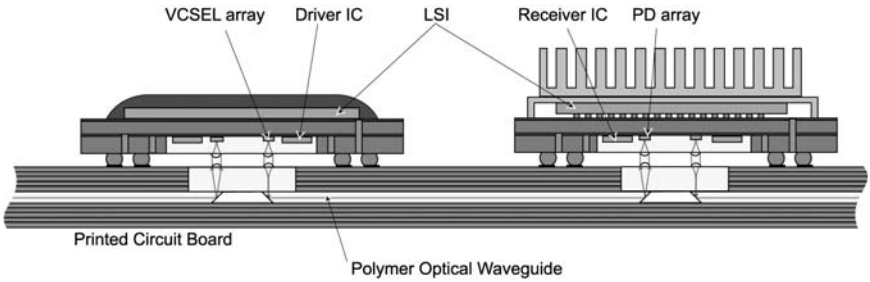


Figure 26.9 Illustration of an interchip interconnect approach based on microlens-assisted coupling to an embedded waveguide layer [26] (© 2003 IEEE).

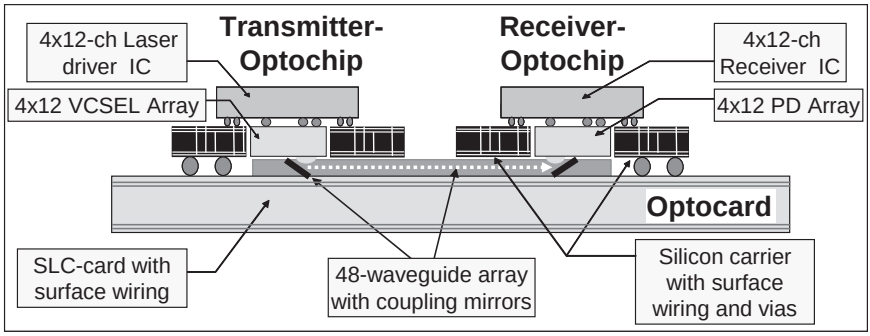


Figure 26.10 Schematic view of the Terabus package [27] (© 2006 IEEE).

Optocard consists of a surface-laminar-circuitry (SLC) with a 150-μm-thick waveguide layer deposited on top. The transmitter and receiver Optochips make extensive use of flip-chip bonding, integrating staggered 4×12 arrays (four rows with individual 250 μm pitch) of 985 nm GaAs-based VCSELs and InP-based pin-type photodiodes (PDs), CMOS driver and receiver ICs, and $10 \times 12 \times 0.3\,\text{mm}^3$ size silicon carrier interposers providing electrical signal routing. These Optochips are then bonded onto the Optocard, all passively aligned. At the chosen operating wavelength, available polymers today have higher attenuation than at 850 nm; here it amounts to 0.1 to 0.16 dB/cm. The main advantage of the longer wavelength is the simplified coupling scheme illustrated in Fig. 26.11. Both the GaAs and InP substrates are transparent, and thus antireflection-coated microlens arrays are etched into the device backsides, which image the active areas onto the $35 \times 35\,\mu\text{m}^2$ size waveguide cores. The 45° slopes of the turning mirrors are made by laser ablation and are gold coated for high reflectivity. The gap between the optoelectronic chips and the waveguide layer is filled by an index-matched optical underfill material. In the paper, no complete link operation has been reported yet; however, transmitter and receiver operation is shown up to 20 and 14 Gbit/s per channel, respectively. This represents the fastest directly modulated VCSEL transmitter with a CMOS driver demonstrated to date. Besides the decrease of the optical link losses, heat extraction during fully parallel operation of the interconnect system is a remaining challenge. To relax the alignment tolerances and enhance the coupling efficiencies, the single-lens coupling in Fig. 26.11 (left) is replaced by a dual-lens relay system in [28]. A full transceiver link is operated at 10 Gbit/s per channel.

Since coupling losses between active elements and the waveguides are an issue in all demonstrators shown before, it is worth mentioning a recently developed option. It consists of flexible polymer pillars that provide waveguiding segments between optoelectronic chips and the mirrors of the board-level waveguides. Figure 26.12 schematically contrasts free-space coupling using a microlens and guided-wave coupling. Using Avtarel polymer pillars with 50-μm diameter and 150-μm height, it is demonstrated that coupling losses from a $30 \times 30\,\mu\text{m}^2$ size

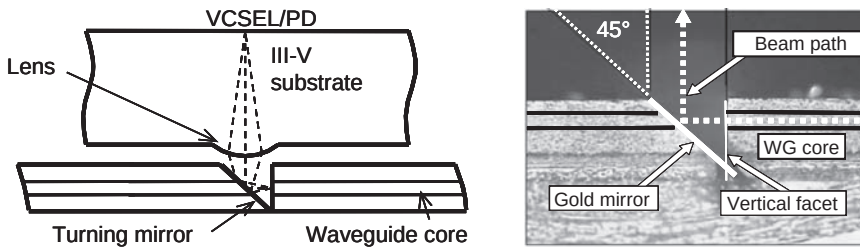


Figure 26.11 Optical coupling scheme between optoelectronic devices and waveguides in the Terabus project (left) and side view of a coupling mirror (right) [27] (© 2006 IEEE).

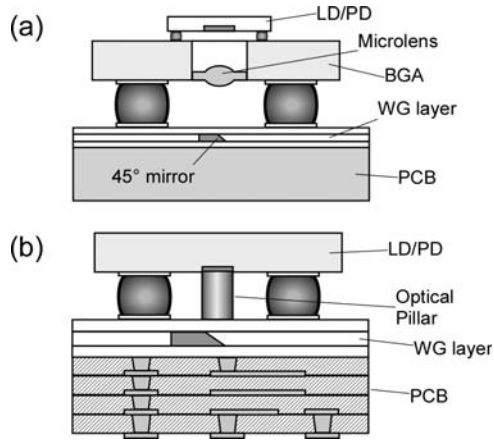


Figure 26.12 Comparison of laser diode (LD) and photodiode (PD) assembly on optical boards using a ball grid array (BGA) package with microlenses (a) and direct active chip attach with light guiding optical pillars (b) [29] (© 2007 IEEE).

waveguide to a chip can be significantly reduced to only 0.5 dB [29]. Moreover, chip-to-substrate lateral displacements (e.g., caused by alignment errors), substrate bow, or coefficient of thermal expansion mismatch can be compensated up to tens of micrometers [30]. More fundamental work on the optical pillars is reported in [31] and shows that mirror angle insensitivity is another advantage of the concept. Up to now, however, it has not been implemented for coupling to a board-embedded waveguide layer.

To this point, all examples in this section have relied on board-level channel waveguides for interchip or intermodule interconnection. We shall now briefly mention a two-dimensional fiber-based approach. Reference [1] discusses fiber options such as ordered fiber bundles (OFBs) or fiber image guides (FIGs; also denoted imaging fiber bundles, IFBs), made either from glass or polymer. An OFB made from 128 step-index polymer optical fibers with core and cladding diameters of 120 and 125 µm, respectively, terminating in a connector containing 90° fiber bends [32], is shown in Fig. 26.13. It is part of a system demonstrator within the early OIIC project of the European Commission [33], where three field-programmable gate array (FPGA) chips are mutually interconnected in order to implement a virtually three-dimensionally stacked processing architecture [34]. The FPGAs contain interleaved analog driver and receiver circuitry, onto which two 4×8-channel VCSEL and photodiode arrays have been flip-chip bonded. The module assembly is all-passive, and no coupling optics are involved. Owing to the available integrated circuit technology generation, the data rate was limited

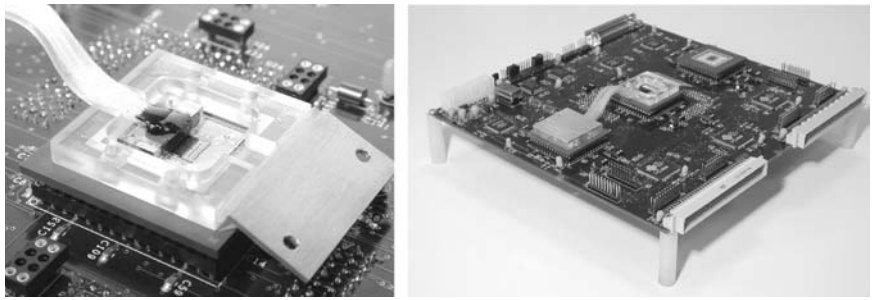


Figure 26.13 Optical coupling unit (left) and entire demonstrator board of the OIIC project of the EC, where plastic optical fiber-based connection is established between two of the three processing chips (right).

to 80 Mbit/s per channel, whereas individual VCSELs were shown to be suited for 12.5 Gbit/s data transmission [35]. More details are contained in [1].

26.4.3. Intrachip Communication

Intrachip optical interconnection is currently a rather speculative field, and it is possible that it will never be implemented. A real need for improved technology is seen at the global on-chip interconnection level. Requirements for optical interconnects such as maximum transmitter and receiver delays are derived from the extrapolated performance of electrical interconnects in [36]. The paper argues that optical solutions cannot compete with their electrical counterparts in terms of bandwidth density on a chip-length scale unless wavelength-division multiplexing is employed. On the other hand, optical interconnects have a delay advantage, but in order to exploit this advantage, they must also have low total power consumption. It is stressed that monolithic integration with the silicon chips is required, where only guided-wave implementations are considered. References to the progress in silicon photonics are made, and it becomes clear that present devices are entirely insufficient. In general, for on-chip optical interconnects, using modulators is considered more realistic than using lasers, owing mainly to temperature and reliability issues. In [37] it is pointed out that the choice between waveguide-based solutions and free-space optical interconnects is a question of the considered class of global interconnects. An arbitrarily configurable free-space scheme using microlenses, microprisms, and a spherically curved mirror is proposed. In [1], several previous free-space optical chip interconnects systems are referenced. In many cases, proof-of-principles have been achieved, and remarkable progress has been made in component integration. However, the optical systems still have a bulky appearance, demanding continued efforts in the fields of microoptics and optomechanics.

A few remarks are also due on the architectural side of the interconnect problem. Two-dimensional optical links are usually characterized by a high spatial regularity of data channels. Thus, it is intuitively clear that electronic processing systems potentially benefit the more from optical chip interconnections, the higher the similarity between the architectures of optical link and electronic chip. In [38] it is emphasized that only appropriate, namely, fine-grained and massively parallel computing architectures, rather than the traditional von Neumann implementation, can exploit the features of short-distance optical interconnections. In particular, this characteristic is met by neural and reconfigurable computing structures. A potential advantage of optical interconnects concerning improved latency over distances even in the mm-range is evidenced in [39] for a special application case, namely, optical fan-out in VLSI systems, where in particular the electrical isolation property of optical interconnects is exploited. As motivated in [40], another strong benefit of on-chip optical interconnects might be experienced in the global clock distribution system, where significant improvements in timing uncertainties (skew and jitter) and total power consumption seem viable. However, both in [39] and [40], practical implementation still fails because compact CMOS-compatible devices like modulators, couplers, or splitters are not available.

To illustrate a free-space optical intrachip interconnect scheme, we include Fig. 26.14 from [41], which is the extension from work described in [42], also contained in [1]. The optical pathway consists of a micro-optical bridge, which is a combination of a glass prism and a plastic baseplate. It integrates functions of beam focusing by microlenses and deflection through total internal reflection. In this case, the electronic chip is traditionally populated with two-dimensional VCSEL and photodiode arrays by hybrid integration. In the design in Fig. 26.14, each channel has an equal optical pathway length of 8 mm, and the interconnect lengths on the chip vary from 385 to 4400 µm. Whereas the given implementation allows only fixed signal routing, a reconfigurable version using beam-splitting diffractive optical elements is considered in [43].

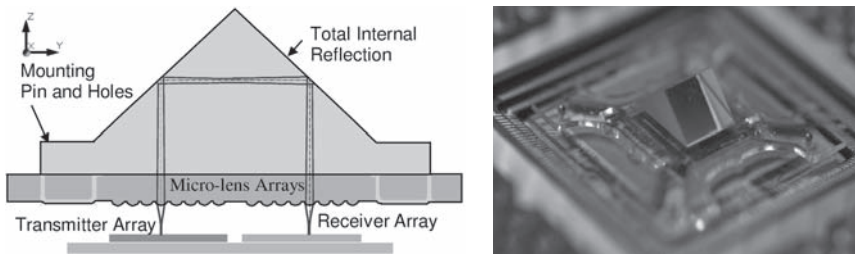


Figure 26.14 Free-space microoptical chip-level interconnection module (left) and prototype of the demonstrator (right) [41] (© 2006 IEEE).

26.5. CONCLUSION

In this chapter we have dealt with ultra-short reach optical interconnects ranging from the backplane down to the intrachip level. In the most advanced demonstrations to date, on hybrid electrical–optical backplanes and boards, optical signals are mainly transported through polymer channel waveguides, which have reached a rather high level of maturity. Material attenuation favors the 850-nm wavelength regime, where high-performance VCSELs are available, featuring properties like low threshold and driving currents, low-divergence symmetric beam profiles, easy one- and two-dimensional array formation, low cost, high reliability, and high-speed modulation capability exceeding 20 Gbit/s. Due to the progress in electronics technologies, optical solutions need at least to run at 10 Gbit/s data rate. VCSELs can meet this target, but attention has to be paid also to the photodetector side, where the device capacitance may limit the speed and may require small waveguide cores, which in turn make alignment and coupling more critical. The most severe obstacles to commercial implementation are on the packaging side. Ideally, the optoelectronic modules should be picked and placed by the same machines as the electronic components, which requires reliable passive self-alignment schemes for coupling to the embedded waveguide layer. Optical connectors between boards and backplane are not yet ready and need to be standardized. Impressive work has been done at the interchip interconnection level; however, the optical layer is still mostly put on the board surface with too much area consumption. Optical interconnects, though holding some promise for the realization of new computing architectures, may never reach the intrachip level owing to the lack of competitive optoelectronic devices. For instance, VCSELs should not be considered here due to very high operating temperatures and postprocessing steps jeopardizing the chip yield. Free-space approaches are discussed at length in the literature but appear to be realistic at best at the chip level. In this field, inventions are urgently required. Of course, these may equally likely be made in the electrical domain, rendering electrical interconnects a constantly moving and improving target for the optics community. Although by far no complete overview over intrabox optical interconnection could be given, this chapter seeks to give a flavor of the variety of approaches existing today. This variety shows both the present weaknesses by not yet having identified the best route and the strengths manifested in the creativity of researchers worldwide.

REFERENCES

1. Michalzik, R. 2002. Optical backplanes, board and chip interconnects. Chapter 6 in *Fiber Optic Data Communication: Technological Trends and Advances*, ed. C. DeCusatis, pp. 216–269. San Diego: Academic Press.
2. Savage, N. 2002, August. Linking with light. *IEEE Spectrum*: 32–36.

3. Uhlig, S., and M. Robertsson. 2006. Limitations to and solutions for optical loss in optical backplanes. *J. Lightwave Technol*. 24, no. 4:1710–1724.
4. Bona, G. L., B. J. Offrein, U. Bapst, C. Berger, R. Beyeler, R. Budd, R. Dangel, L. Dellmann, and F. Horst. 2004. Characterization of parallel optical-interconnect waveguides integrated on a printed circuit board. *Proc. SPIE*, 5453:134–141.
5. Naeemi, A., J. Xu, A. V. Mule, T. K. Gaylord, and J. D. Meindl. 2004. Optical and electrical interconnect partition length based on chip-to-chip bandwidth maximization. *IEEE Photon. Technol. Lett*. 16, no. 4:1221–1223.
6. Cho, H., P. Kapur, and K. C. Saraswat. 2004. Power comparison between high-speed electrical and optical interconnects for interchip communication. *J. Lightwave Technol*. 22, no. 9:2021–2033.
7. Moisel, J., J. Guttmann, H.-P. Huber, O. Krumpholz, M. Rode, R. Bogenberger, and K.-P. Kuhn. 2000. Optical backplanes with integrated polymer waveguides. *Opt. Eng*. 39, no. 3:673–679.
8. Yoon, K. B., I.-K. Cho, S. H. Ahn, M. Y. Jeong, D. J. Lee, Y. U. Heo, B. S. Rho, H.-H. Park, and B.-H. Rhee. 2004. Optical backplane system using waveguide-embedded PCBs and optical slots. *J. Lightwave Technol*. 22, no. 9:2119–2127.
9. Eldada, L., and L. W. Shacklette. 2000. Advances in polymer integrated optics. *IEEE J. Select. Topics Quantum Electron*. 6, no. 1:54–68.
10. Glebov, A. L., J. Roman, M. G. Lee, and K. Yokouchi. 2005. Optical interconnect modules with fully integrated reflector mirrors. *IEEE Photon. Technol. Lett.* 17, no. 7:1540–1542.
11. Hendrickx, N., J. Van Erps, G. Van Steenberge, H. Thienpont, and P. Van Daele. 2007. Laser ablated micromirrors for printed circuit board integrated optical interconnections. *IEEE Photon. Technol. Lett.* 19, no. 11:822–824.
12. Chen, R. T., L. Lin, C. Choi, Y. J. Liu, B. Bihari, L. Wu, S. Tang, R. Wickmann, B. Picor, M. K. Hibbs-Brenner, J. Bristow, and Y. S. Liu. 2000. Fully embedded board-level guided-wave optoelectronic interconnects. *Proc. IEEE* 88, no. 6:780–793.
13. Glebov, A. L., M. G. Lee, S. Aoki, D. Kudzuma, J. Roman, M. Peters, L. Huang, D. S. Zhou, and K. Yokouchi. 2006. Integrated waveguide microoptic elements for 3D routing in board-level optical interconnects. *Proc. SPIE* 6126:61260N-1–11.
14. Rho, B. S., S. Kang, H. S. Cho, H.-H. Park, S.-W. Ha, and B.-H. Rhee. 2004. PCB-compatible optical interconnection using 45°-ended connection rods and via-holed waveguides. *J. Lightwave Technol*. 22, no. 9:2128–2134.
15. Hwang, S. H., M. H. Cho, S.-K. Kang, T.-W. Lee, H.-H. Park, and B. S. Rho. 2007. Two-dimensional optical interconnection based on two-layered optical printed circuit board. *IEEE Photon. Technol. Lett.* 19, no. 6:411–413.
16. Cho, H. S., S. Kang, B. S. Rho, H.-H. Park, K.-U. Shin, S.-W. Ha, B.-H. Rhee, D.-S. Ku, S. T. Jung, and T. Kim. 2004. Fabrication of fiber-embedded boards using grooving technique for optical interconnection applications. *Opt. Eng*. 43, no. 12:3083–3088.
17. Choi, J. H., L. Wang, H. Bi, and R. T. Chen. 2006. Effects of thermal-via structures on thin-film VCSELs for fully embedded board-level optical interconnection system. *IEEE J. Select. Topics Quantum Electron*. 12, no. 5:1060–1065.
18. Choi, C., L. Lin, Y. Liu, and R. T. Chen. 2003. Performance analysis of 10-μm-thick VCSEL array in fully embedded board level guided-wave optoelectronic interconnects. *J. Lightwave Technol.* 21, no. 6:1531–1535.
19. Hiruma, K., M. Kinoshita, S. Hiramatsu, and T. Mikawa. 2004. Epitaxial lift-off of GaAs/AlGaAs films with vertical cavity surface emitting lasers for high-density packaging of optoelectronic interconnections. *Jap. J. Appl. Phys*. 43, no. 10:7054–7057.
20. Cho, S.-Y., S.-W. Seo, N. M. Jokerst, and M. A. Brooke. 2004. Board-level optical interconnection and signal distribution using embedded thin-film optoelectronic devices. *J. Lightwave Technol*. 22, no. 9:2111–2118.

21. International Technology Roadmap for Semiconductors; see URL http://www.itrs.net/reports.html.
22. Miller, D. A. B. 2005. Opportunities for optics to silicon chips. Proc. *IEEE Lasers and Electro-Opt. Soc. Ann. Meet., LEOS 2005*:202–203. Sydney, Australia.
23. Zheng, X., J. K. Lexau, J. Bergey, J. E. Cunningham, R. Ho, R. Drost, and A. V. Krishnamoorthy. 2007. Optical transceiver chips based on co-integration of capacitively coupled proximity interconnects and VCSELs. *IEEE Photon. Technol. Lett*. 19, no. 7:453–455.
24. Collet, J. H., D. Litaize, J. Van Campenhout, C. Jesshope, M. Desmulliez, H. Thienpont, J. Goodman, and A. Louri. 2000. Architectural approach to the role of optics in monoprocessor and multiprocessor machines. *Appl. Opt*. 39, no. 5:671–682.
25. Hiramatsu, S., and T. Mikawa. 2006. Optical design of active interposer for high-speed chip level optical interconnects. *J. Lightwave Technol*. 24, no. 2:927–934.
26. Ishii, Y., N. Tanaka, T. Sakamoto, and H. Takahara. 2003. Fully SMT-compatible optical-I/O package with microlens array interface. *J. Lightwave Technol*. 21, no. 1:275–280.
27. Schares, L., J. A. Kash, F. E. Doany, C. L. Schow, C. Schuster, D. M. Kuchta, P. K. Pepeljugoski, J. M. Trewhella, C. W. Baks, R. A. John, L. Shan, Y. H. Kwark, R. A. Budd, P. Chiniwalla, F. R. Libsch, J. Rosner, C. K. Tsang, C. S. Patel, J. D. Schaub, R. Dangel, F. Horst, B. J. Offrein, D. Kucharski, D. Guckenberger, S. Hegde, H. Nyikal, C.-K. Lin, A. Tandon, G. R. Trott, M. Nystrom, D. P. Bour, M. R. T. Tan, and D. W. Dolfi. 2006. Terabus: Terabit/second-class card-level optical interconnect technologies. *IEEE J. Select. Topics Quantum Electron*. 12, no. 5:1032–1044.
28. Doany, F. E., C. L. Schow, C. Baks, R. Budd, Y.-J. Chang, P. Pepeljugoski, L. Schares, D. Kuchta, R. John, J. A. Kash, F. Libsch, R. Dangel, F. Horst, and B. J. Offrein. 2007, May/June. 160-Gb/s bidirectional parallel optical transceiver module for board-level interconnects using a single-chip CMOS IC. Proc. *57th Electronic Technology and Components Conference, ECTC 2007*: 1256–1261. Reno, Nev.
29. Glebov, A. L., C. J. Uchibori, and M. G. Lee. 2007. Direct attach of photonic components on substrates with optical interconnects. *IEEE Photon. Technol. Lett*. 19, no. 8:547–549.
30. Glebov, A. L., D. Bhusari, P. Kohl, M. S. Bakir, J. D. Meindl, and M. G. Lee. 2006. Flexible pillars for displacement compensation in optical chip assembly. *IEEE Photon. Technol. Lett*. 18, no. 8:974–976.
31. Ogunsola, O. O., H. D. Thacker, B. L. Bachim, M. S. Bakir, J. Pikarsky, T. K. Gaylord, and J. D. Meindl. 2006. Chip-level waveguide-mirror-pillar optical interconnect structure. *IEEE Photon. Technol. Lett*. 18. no. 15:1672–1674.
32. Neyer, A., B. Wittmann, and M. Jöhnck. 1999. Plastic-optical-fiber-based parallel optical interconnects. *IEEE J. Select. Topics Quantum Electron*. 5, no. 2:193–200.
33. OIIC stands for "optically interconnected integrated circuits"; see URL http://oiic.intec.ugent.be/index.htm.
34. Brunfaut, M., W. Meeus, J. Van Campenhout, R. Annen, P. Zenklusen, H. Melchior, R. Bockstaele, L. Vanwassenhove, J. Hall, B. Wittmann, A. Neyer, P. Heremans, J. Van Koetsem, R. King, H. Thienpont, and R. Baets. 2001. Demonstrating optoelectronic interconnect in a FPGA based prototype system using flip chip mounted 2D arrays of optical components and 2D POF-ribbon arrays as optical pathways. *Proc. SPIE* 4455:160–171.
35. King, R., R. Michalzik, D. Wiedenmann, R. Jäger, P. Schnitzer, T. Knödl, and K. J. Ebeling. 1999. 2D VCSEL arrays for chip-level optical interconnects. *Proc. SPIE* 3632:363–372.
36. Haurylau, M., G. Chen, H. Chen, J. Zhang, N. A. Nelson, D. H. Albonesi, E. G. Friedman, and P. M. Fauchet. 2006. On-chip optical interconnect roadmap: challenges and critical directions. *IEEE J. Select. Topics Quantum Electron*. 12, no. 6:1699–1705.
37. Haney, M. W., M. Iqbal, and M. J. McFadden. 2005. Optical interconnects for intrachip global communication: motivation & validation. Proc. *IEEE Lasers and Electro-Opt. Soc. Ann. Meet., LEOS 2005*:206–207. Sydney, Australia.

38. Fey, D., W. Erhard, M. Gruber, J. Jahns, H. Bartelt, G. Grimm, L. Hoppe, and S. Sinzinger. 2000. Optical interconnects for neural and reconfigurable VLSI architectures. *Proc. IEEE* 88, no. 6:838–848.
39. Pappu, A. M., and A. B. Apsel. 2006. Demonstration of latency reduction in ultra-short distance interconnections using optical fanout. *IEEE J. Select. Topics Quantum Electron*. 12, no. 6:1664–1670.
40. Cassan, E., D. Marris, M. Rouvière, L. Vivien, and S. Laval. 2005. Comparison between electrical and optical global clock distributions for CMOS integrated circuits. *Opt. Eng*. 44, no. 10:105402-1–105402-10.
41. Vervaeke, M., C. Debaes, B. Volckaerts, and H. Thienpont. 2006. Optomechanical Monte Carlo tolerancing study of a packaged free-space intra-MCM optical interconnect system. *IEEE J. Select. Topics Quantum Electron*. 12, no. 5:988–996.
42. Thienpont, H., C. Debaes, V. Baukens, H. Ottevaere, P. Vynck, P. Tuteleers, G. Verschaffelt, B. Volckaerts, A. Hermane, and M. Hanney. 2000. Plastic microoptical interconnection modules for parallel free-space inter- and intra-MCM data communication. *Proc. IEEE* 88, no. 6:769–779.
43. Artundo, I., L. Desmet, W. Heirman, C. Debaes, J. Dambre, J. M. Van Campenhout, and H. Thienpont. 2006. Selective optical broadcast component for reconfigurable multiprocessor interconnects. *IEEE J. Select. Topics Quantum Electron*. 12, no. 4:828–837.

27

Silicon Photonics

Nahum Izhaky
Intel Corporation, Fab8, Jerusalem, Israel

27.1. INTRODUCTION

27.1.1. What is Silicon Photonics? Pros and Cons of Si for Photonics

Integrated photonics is a main candidate as photonic technology to provide the various demands of numerous fields such as communication, computing, imaging, and sensing. It refers to miniaturized optoelectronic systems, that is, to combining discrete very small optical and electronic elements on a common substrate by means of optical waveguides (WGs) and metal lines. Many material systems were proposed as a platform for such a solution, such as $LiNbO_3$, InP, GaAs, Silica on Silicon, polymers, and glass. Some of them have already succeeded in providing commercial devices (e.g., modulators in $LiNbO_3$ or AWGs in silica on silicon), but all suffer from being too expensive, based on nonmature fabrication processes, and far from allowing monolithic integration with conventional electronic circuits.

Silicon has been the mainstay of the electronics industry for the last 40 years and has revolutionized the way the world operates. Today a silicon chip the size of a fingernail contains nearly one billion transistors and has the computing power that only a decade ago would take up an entire room full of servers. In addition the silicon integrated chip has moved beyond just computer microprocessors to impact everything from cell phones to music players to cameras.

Silicon photonics is the emerging technology of producing optical devices and circuits using silicon as the core material with standard CMOS (complementary metal oxide semiconductor) manufacturing equipment and processes. Silicon photonics, based mainly on silicon on insulator (SOI), has recently attracted a great deal of attention. The intrinsic bandgap of silicon (1.1 eV) means that silicon

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

is transparent at wavelengths typically used for optical communication transmission (i.e., 1270 nm to 1625 nm). This allows one to build optical devices in silicon that can route, direct and manipulate light. The utilization of standard CMOS compatible process for silicon photonics allows an incomparable monolithic integration of photonics devices with electronic control functions that offer an opportunity for low-cost optoelectronic solutions. In fact, the large refractive index contrast $\Delta n \approx 2$ ($n_{core} \approx 3.5$, $n_{clad} \approx 1.5$, $\Delta n \equiv n_{core} - n_{clad}$) of the silicon/oxide waveguide system makes it favorable to fabricate high-density photonic circuits. Other optical materials used in integrated optics provide lower index contrast of about 2–3 orders of magnitude; hence, Si photonics allows reducing the minimal waveguide bend radius from centimeters/millimeters to micrometers, thereby yielding higher area condensation of about 2–3 orders of magnitude (very small footprints) and reduced power consumption.

There are several reasons for silicon photonics, low-cost solutions. (1) Si and SOI bare wafers cost per area is the lowest in the market in relation to other materials. (2) CMOS processing provides Si and SOI processed wafers with the best available line yield and die yield—most mature technology. (3) Most CMOS Fabs work under high load and provide high-volume manufacturing (HVM) products. This allows much lower processing activity cost than with low loading Fabs. (4) The above high indexes contrast (miniaturization), together with large Si wafers area (150–300 mm diameter), further reduce the cost per chip. (5) Some monolithic optoelectronic integration of optical and electronic elements may shrink dimensions further and also reduce total cost (e.g., integration of photodetectors with their TIA amplifiers), altogether allowing 10–100× cost and price reduction!

Until 2004 it was largely believed that silicon photonic devices were practically limited in bandwidth to about several tens of MHz. This drawback was entirely removed when 10–40 Gb/s devices with several physical mechanisms were demonstrated (e.g., free carrier plasma dispersion effect in MOS capacitor [1], and reverse biased PN junctions [2]) in various structures (e.g., Mach-Zehnder interferometers—MZIs and Ring resonators).

Most communication semiconductor lasers operate in the near-infrared (NIR) wavelength (around 0.85, 1.31, and 1.55 μm), a region where silicon is a poor detector. In order to improve the performance of silicon-based detectors, the most common approach is to grow germanium (Ge) on the Si and to use the Ge narrower bandgap to extend the maximum detectable wavelength. Today various groups have already reported on Ge-on-Si (NIR) detectors, with responsivity nearly equivalent to those of III–V based devices [3].

The Si high indexes contrast comes with a problem. The optical modes of a Si waveguide are significantly different from those in conventional optical fibers. Hence, a transition between two such elements (one from the Si chip and the other from the outer world) without any special care may create significant signal

loss of 5–20 dB/facet. Various mode converter solutions have been demonstrated [4–6] to reduce this loss to a reasonable level of 1–1.5 dB/facet.

Silicon's main drawback is its poor optical emission efficiency due to its indirect band-gap. Bands are separated not just by energy but also by momentum—photon emission requires the participation of phonons, hence, orders of magnitude lower radiative emission probability than with direct semiconductors. Although various techniques and vast efforts have been made to invent efficient and reliable pure silicon electrically pumped laser, none has succeeded and an electrically pumped monolithic silicon laser is still missing [7]. Nevertheless, significant progress has been made in the field of hybrid silicon laser [8] to allow low-cost and efficient laser light that looks promising in regard to future optical interconnects requirements and other applications. In addition, the Raman gain in silicon waveguides is ~5 orders of magnitude stronger than in silica optical fibers. This fact was employed to demonstrate very efficient and compact pure silicon Raman lasers (optically pumped) in both pulse and CW modes [9–11]. Raman silicon lasers open additional markets for the silicon photonics technology.

One of the main limiting obstacles preventing optics from further penetration into new domains is the need for very low-power consumption elements and circuits. Many applications such as o-e-o (optical-electrical-optical) signal conversions, transceiver arrays, fast interconnect, and routing/switching, require more energy-efficient solutions than optical devices offer currently. However, the high miniaturization (nanoscale) potential silicon photonics, together with the efficient refractive index fast tuning, show that these low-power elements and circuits are attainable [12].

Silicon photonic waveguide components propose even more far-reaching bandwidth than many other materials. Its wavelength operation window can be extended beyond the optical telecom range (1.2–1.6 μm) even up to 100 μm. This ultra-high bandwidth of Si waveguides opens numerous future applications for silicon photonics.

If we combine the above main merits into one figure of merit (FOM) for silicon photonics to be defined $FOM \equiv (speed \times reliability)/(power \times area \times cost)$, then the expected overall benefit from silicon photonics can be emphasized quantitatively. For example, comparing the estimated FOM of optical interconnects based on silicon photonics with current copper interconnects, although some parts are inconclusive, makes it easy to estimate an improvement of several orders of magnitude in favor of silicon photonics with no degradation in any of the above parameters.

The main issues that still need further development to fulfill silicon photonics' high potential are first to optimize the elements' performance; second, to develop integrated optoelectronic circuit solutions; and last but not least, to further develop the maturity of this just emerging technology in all facets—design, testing, processing standards, packaging, reliability, and market. Recent research advances

in silicon photonic devices performance have come at a rapid pace. However, significant effort will be needed before one can commercialize this technology. One of the keys to successful commercialization will be developing a high-volume manufacturing (HVM) process that will enable silicon photonic devices to be processed in a CMOS fabrication facility. Furthermore, these photonic devices must be processed alongside existing CMOS electronic wafers in order to amortize their costs. This will put significant constraints on the processes to be developed and the technologies implemented to manufacture these silicon photonic devices.

27.1.2. Silicon Photonics Applications

Based on its many benefits, silicon photonics shows numerous potential applications within the following fields:

- **Discrete optical elements and modules for optical communications** (Long haul, Metro, and Access). For example, optical modulators, photodetectors, variable optical attenuators (VOAs), and the most desirable—miniature transceivers that can revolutionize the whole field of photonics in communication, computing, imaging, and allowing real systems on a chip (SOC). The use of silicon as base material allows, for the first time, mass production of these elements. The main focus in silicon photonics in the last five years was to prove the concept of these building blocks and to improve their performance. These devices are described later in this chapter.
- **Optical interconnects.** Currently the main application of silicon photonics is for optical interconnects replacing copper to allow higher bandwidth and long reach for faster and larger data centers, rack-to-rack, board-to-board, and chip-to-chip interconnects. It may allow faster computing and enlarged memory in future computers. This is definitely the jewel in the crown among all currently recognized silicon photonics applications with high-volume manufacturing (HVM) demands.
- **Optical sensors and Lab on a chip applications.** Silicon photonics starts to show benefits in various sensing systems such as physical, chemical, bio, and the more integrated lab on a chip. Fabrication of miniature optical-biological-chemical labs on a single silicon die with tiny reactors and microsensors promise a huge advantage over conventional testing methodologies. It requires smaller sample volumes and fewer materials, and it provides faster response time and lower cost.
- **Radio over fibers (RoF).** RoF is an emerging area where silicon photonics could be used for further distributing and transmitting wireless connectivity.
- **Nonlinear optics applications—wavelength conversion, lasers, and optical amplifiers.** Silicon demonstrates enhanced nonlinear effects compared

to other materials (e.g., optical fibers). In addition, its higher index contrast further increases the nonlinear effect (higher intensities). Hence, compact low-cost devices are expected for the benefit of various fields requiring guided amplifiers, lasers, and wavelength converters. Some of these devices will be covered in subsection 27.3.4.

27.2. PASSIVE SILICON PHOTONIC DEVICES

Complicated photonic integrated circuits are usually made of various fundamental photonic building blocks or components. In this section, we describe some selected passive elements and devices such as straight and curved waveguide, directional coupler, multimode interference coupler, Y junction, Mach-Zehnder interferometer, Bragg grating, and ring resonator. In passive devices in addition to the stricter definition of elements with no gain, we mean also elements without any modification in refractive index, that is, no tuning effect such as thermal or plasma dispersion.

27.2.1. Straight and Curved Waveguides

The waveguide is the most fundamental building block from which every guided wave device is built. Interest in silicon-based planar lightwave circuits using SOI platform has been growing for over a decade. SOI wafers, originally developed for integrated electronic circuits, are an ideal substrate for photonic applications, as the buried oxide (BOX–SiO_2) serves well as a bottom cladding that can eliminate any leakage of the lightwave signal into the substrate. Silicon photonics employ two main types of waveguides; the first is of a rib profile (see Fig. 27.1), and the second is of a strip (channel) waveguide where the Si epi layer is being etched to the box. Single-mode strip silicon waveguides are very small

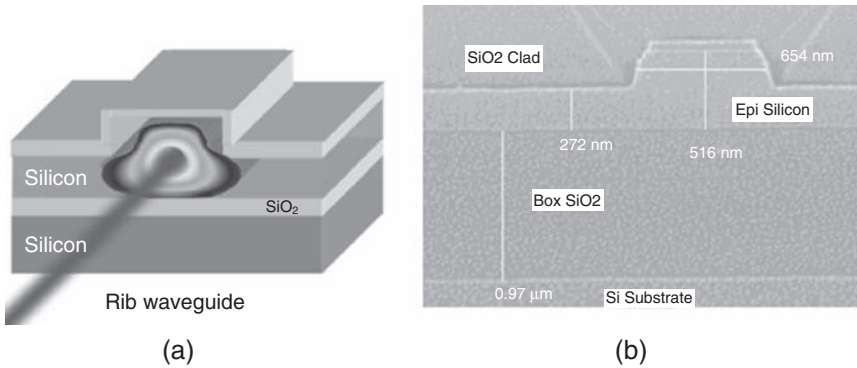


Figure 27.1 (a) Schematic illustration of rib WG cross section and its optical propagating mode. (b) SEM cross section picture of a small rib waveguide (W = 0.65 μm H = 0.5 μm).

(~0.25 μm^2), and a good waveguide is difficult to achieve due to its relatively high index contrast. First, coupling to standard fibers provides very high loss (10–20 dB/facet); second, a very good control on sidewall roughness is required (more than with rib waveguide) for low propagation loss; and third, a relatively high lithography resolution is required for these dimensions and its roughness restrictions. For single-mode rib waveguides, a larger cross section is possible, which significantly improves performance and requires less complicated tools. However, strip waveguides allow bend radius on the order of micrometers, whereas ribs permit bends on the order of hundreds of micrometers. Nevertheless, numerous applications can suffice with the rib, which already provides 1–2 orders of magnitude reduction in radius compared to other technologies ($LiNbO_3$, silica on silicon, polymers) and brings significantly smaller footprint devices and circuits.

Forming low-loss silicon WGs requires reduction of the roughness in the core-cladding interface. This is becoming increasingly important for large-scale integration as the WG dimensions are reduced [13]. Modeling shows that a roughness root mean square (RMS) of less than 3 nm should be achieved to maintain reasonable scattering loss for rib waveguide size of less than 1–2 μm.

In some cases, obtaining low loss in the small-size rib waveguide is critical for achieving a functional device as for Raman net gain and lasing in silicon [10]. This was achieved mainly through the following processing steps: prior to rib patterning, a thin layer of thermal oxide serving as a hard-mask was grown on a low doped (10^{15}/cm^3) SOI wafer of 1-μm BOX. Standard photolithography was used, followed by hard-mask reactive ion etch (RIE) and photoresist clean. The silicon plasma etch conditions are the dominant loss contributors, as the final silicon surface for rib/ridge waveguide is highly sensitive to the gas flow ratio and other tool parameters. Following a careful design of experiment of the Si etch conditions, we have achieved silicon bottom surface roughness of ~0.5 nm as measured by AFM and sidewall RMS roughness of ~2 nm. This process yields ~0.2 dB/cm for waveguide with effective size of 1.7 μm^2 [1].

For smaller waveguide size (W = 0.65, H = 0.5) an additional smoothing oxidation 70 nm was utilized for further reduction of the silicon lithography and dry-etch-related roughness. The effect of oxidation was measured to improve loss from ~2.5 dB/cm to ~0.8 dB/cm [see Fig. 27.1(b)].

Next to the straight waveguides, the second most basic element is the bent waveguide. It is typically used for propagation redirection of the optical mode, which is an indispensable functionality in many optical circuits. Figure 27.2(a) shows schematically a bent waveguide in SOI with a rib width of *W*, a rib height of *H*, and an etch depth of *h*. Transverse optical mode is confined in the xy-plane and propagates in the xz-plane. The inner bend radius of the waveguide is R_1, and the outer radius R_2 (the waveguide width is therefore $W = R_2 - R_1$). For practical photonic circuits, the bent waveguide is usually combined with straight

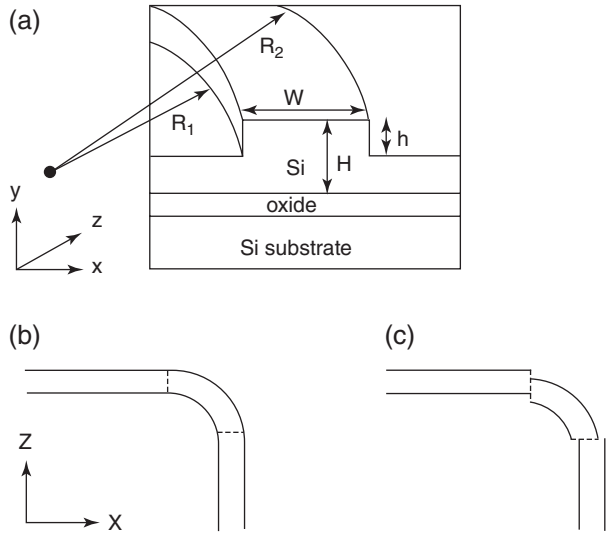


Figure 27.2 (a) Schematic of a waveguide bend in SOI, (b) top view of a 90° bend without lateral offset, and (c) top view of a 90° bend with lateral offset.

waveguides to achieve the device functionalities. As an example, Fig. 27.2(b) and (c) show a top view of a 90° bend waveguide section placed in between two straight waveguides with [Fig. 27.2(c)] and without [Fig. 27.2(b)] a lateral offset.

Physically, bending of a waveguide leads to optical loss as the mode propagates around the bend because of radiation loss from its modal field in the cladding. This radiation loss is inherent to the waveguide bend, which is different from the scattering loss of a straight waveguide due to the waveguide roughness. However, a combination of curved waveguide with roughness in its outer sidewall might increase the radiation loss with larger (than straight waveguide) scattering loss, due to a stronger overlap between the optical mode and the rough side wall. In addition, there is modal abrupt transition loss between the bent waveguide and straight waveguide due to the modal profile and location mismatch. The magnitude of the bend loss depends on the bend radius and the waveguide confinement in the bend direction. For SOI waveguide, the waveguide bend usually occurs in the plane parallel to the silicon-buried oxide interface (xz-plane in Fig. 27.2) due to the device fabrication; therefore, the waveguide bend has most impact on the TE mode.

The mode of the waveguide bend tends to shift to the outer edge of the bend: the smaller the bend radius, the larger the modal shift. The mode distortion in a rib waveguide occurs mainly in the slab region. Because the bent waveguide mode is shifted toward the outer edge of the bend, one could reduce the mode mismatch

induced transition loss by introducing a lateral shift between straight waveguide and bent waveguide [14], as shown in Fig. 27.2(c). The minimum bend radius that results in negligible bend related loss, including radiation loss and straight-to-bend waveguide transition loss, is an important design parameter in integrated optics. The size of the waveguide bend ultimately determines the maximum density with which photonic circuits can be integrated on a single chip. It is well known that the silicon strip (channel) waveguide has a very small bend radius (a few micrometers) because of the high refractive index contrast and waveguide geometry. However, the single-mode operation requires a subwavelength waveguide size, which leads to a relatively high transmission loss (~3 dB/cm) due to fabrication-induced waveguide imperfections [15]. In contrast, the silicon rib waveguide exhibits single-mode behavior for a large range of waveguide sizes. The waveguide transmission loss can be much smaller. Because the radiation loss depends on the lateral confinement strength of the waveguide, for silicon rib waveguide, the minimum bend radius is strongly dependent on the waveguide geometry. The bend radius is smaller for smaller waveguide cross sections. The bend radius can be as small as ~50 μm for a rib waveguide with rib height of 0.5 um. This makes the small cross section rib waveguide very attractive for both active and passive devices. One can also minimize the bend loss by using a deeper rib etching on the outer edge of the bend. For example, low loss of an asymmetric rib waveguide bend has been reported [16].

27.2.2. Fiber Waveguide Coupling

Optical integrated circuits need to be connected to the outer world. However, trivial butt coupling between standard optical fibers to micron or submicron silicon waveguides produces significant unacceptable loss (of 5–20 dB/facet) due to mode mismatch. Therefore, it is necessary to develop special elements that couple between silicon waveguides and optical fibers.

Silicon-single mode waveguides with their relatively very high refractive index and very small dimensions present even higher losses than other material systems when coupled to optical fibers. In addition, significant different refractive indexes for the silicon and silica produce considerable Fresnel's reflections ($[(3.5 - 1.5)/(3.5 + 1.5)]^2 = 16\%$) of ~1.5 dB loss for two facets, and a backreflection of about −8 dB/facet occurs. Hence, an antireflection coating is necessary to eliminate the above detrimental effects. Besides the natural mode mismatch and the Fresnel's reflection one should develop an accurate (negligible misalignments), fast, and low-cost fiber pigtailing process. This coupling should also be wavelength insensitive and polarization independent. This will be a significant engineering development in the future evolution of silicon photonics.

In order to couple between fiber and a waveguide, one should match between their optical modal field profiles. The power coupling efficiency between two

different modes (η) is given by the overlap integral between modes E_1 and E_2 of the waveguide and the fiber

$$\eta = \frac{\left|\iint E_1^*(x, y)E_2(x, y)dxdy\right|^2}{\iint|E_1|^2dxdy\iint|E_2|^2dxdy} \tag{27.1}$$

For example, a rib with the dimensions W = 1.5 μm, H = 1.5 μm, and h = 0.75 μm. (MFD ≈ 1.5 μm) produce coupling efficiency of ~10 dB/facet with a SMF (MFD ≈ 9 μm). It is obvious that a mode converter is needed.

Various methods have been suggested and tested to obtain such a converter. The most obvious design is the 3-D lateral and vertical adiabatic taper. The vertical tapering requires differential etch depth along the taper length with very small (scale of nanometers) roughness. Such a solution requires nonstandard-complicated and expensive processing especially for coupling, with very small rib or strip waveguides. Hence, other techniques were developed. Bookham Technology developed a nonvertical taper as shown in Fig. 27.3 [6, 17, 18]. This compact design allows good (~0.5 dB/facet) butt coupling for cross sections of 2 μm and above and requires the assistance of a lensed fiber for smaller dimensions. The operation principle is based on two stages. First, lightwave is butt coupled to a mode-matched Si interface and then, due to the adiabatic lateral tapering, the effective index is adiabatically reduced and light is more and more contained within the waveguide until a negligible part of it is within the upper taper (light is no longer guided in the too narrow upper taper). A different solution introduces the assistance of waveguide gratings [4–6]. Although these grating solutions theoretically predict very low coupling loss, practically this is very hard to achieve and is naturally less attractive due to its polarization dependence. Based on the numerous suggested methods, it is clear that this is an important task and

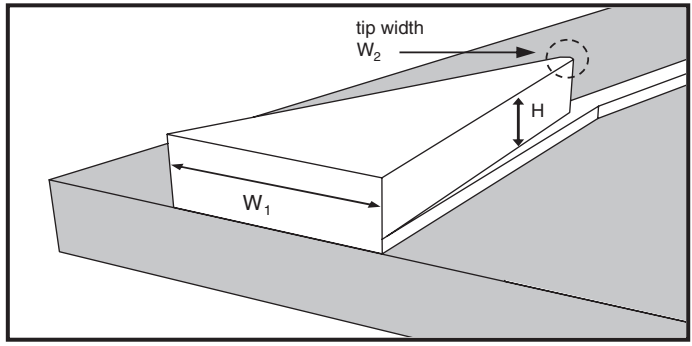


Figure 27.3 The SSC is built from 2 levels: (1) Rib WG; (2) taper with opening width of W1, final width W2 and height H.

not easily accomplished. Another suggestion was made by Doylend and Knights [19] to separate the upper taper (as in the Bookham proposal) and the waveguide by a thin oxide layer. As a result, a vertical directional coupler couples the inserted matched light into the waveguide. Doylend and Knights obtained theoretical estimation of less than 0.5 dB/facet. Still, this has to be proved practically, and one probably encounters polarization and wavelength sensitivities. Luxtera has suggested an "out of the box" solution of normal incidence fiber to waveguide coupling via a suitable surface curved grating geometry [20]. This technique is a great step forward toward automatic wafer level low cost and fast testing. However, it is naturally both wavelength sensitive and polarization dependent. An additional potential solution especially for strip waveguides is the inverted taper [5, 6]. With this design the edge of the waveguide is narrowed significantly (~100 nm) and adiabatically, thus allowing mode matching based on a non-confined mode effect. This technique, though very compact and attractive, requires high-resolution lithography and appropriate bulk cladding. In Intel we develop and test various techniques.

The dual taper is a generalization of the 2-D mode converter taper [17, 20], which has very low polarization dependence and CMOS compatible processing, and can be easily integrated with all silicon-based devices active and passive. This generalization allows fiber waveguide coupling even to small cross-sectional waveguides (>0.5 μm). The end facet of the chip provides larger end facets and better matches the modes. The taper design consists of a double stage taper, DST, with three processing steps: (1) bottom taper (BT); (2) WG definition; and (3) upper taper (UT) (see Fig. 27.4). Results are very promising: polarization independent 1.5 dB/facet total taper loss, using standard SMF. Moreover, the taper performance is preserved over a wide wavelength range (1.31, 1.55 μm).

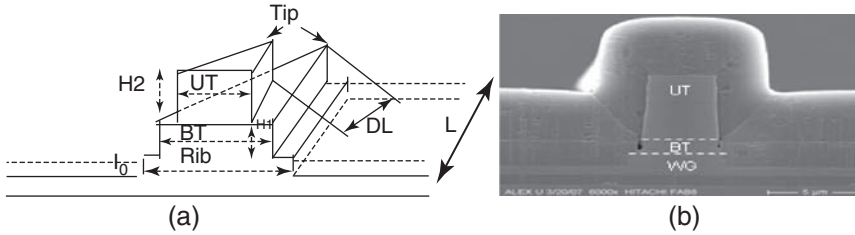


Figure 27.4 (a) Schematics of a DST, lower part 12 μm width is the WG input; BT is the lower taper with input of 12 μm and UT with the input of 11 μm. ΔL = 150 μm represent the spacing between the tip of the BT to UT. Where h = 1.5 μm, H1 = 2.5 μm, and H2 = 5.5 μm. (b) SEM cross section of the DST.

27.2.3. Couplers and Splitters

27.2.3.1. Directional Coupler

Directional coupler (DC) is one of the fundamental and extensively used elements in integrated optics and optical communications. In its most basic configuration, a DC consists of two single-mode waveguides (WGs) in close proximity with input/output waveguide fan-out, as illustrated in Fig. 27.5(a). Its main function is to exchange optical power between the two adjacent WGs due to the modal interaction. A general refractive index profile of the two coupled WGs and its realization using SOI rib waveguides are shown in Fig. 27.5(b) and (c).

When the WG separation, s, is small enough, the evanescent parts of the guided modes overlap, and coupling between the two WGs occurs. In the case where the coupling is weak, the proximity of one WG to the other can be considered a perturbation. Thus, a combination of the evanescent tail of the mode guided in one WG and the neighboring WG induces an electric perturbing polarization generating a wave in the neighboring WG, which in turn may couple back to the first WG. One expects stronger coupling for weaker mode confinement and better overlap. Therefore, narrower WGs, longer wavelengths, or smaller WG separation

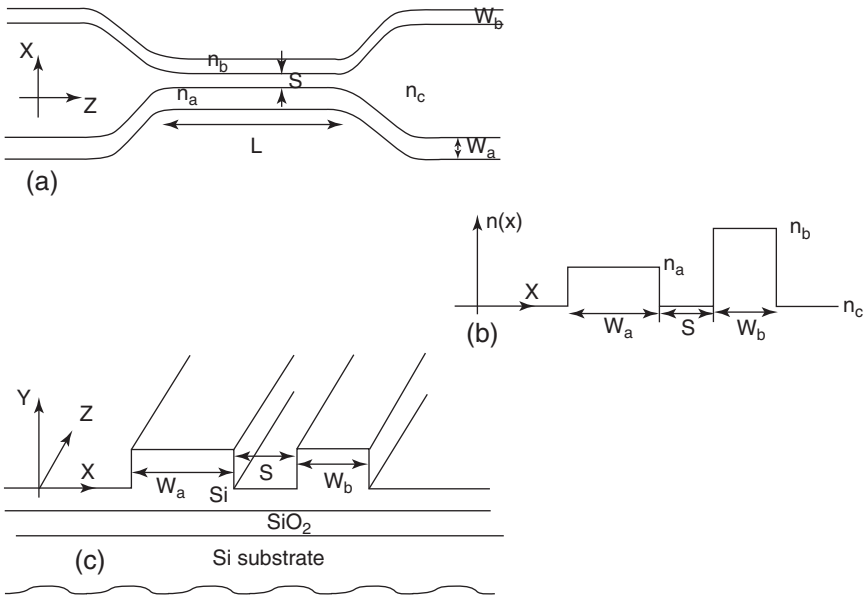


Figure 27.5 (a) Top view of a directional coupler, (b) refractive index distribution, and (c) realization in SOI rib waveguides.

(gap) provides stronger coupling and hence more compact elements. In addition, a pair of symmetric WGs is necessary for full power transfer, whereas asymmetric WGs (e.g., $W_a \neq W_b$, or $n_a \neq n_b$) cannot achieve full coupling. Overall, as in other coupled systems, the lightwave is codirectionally coupled back and forth between the two WGs in a periodic manner, in accordance with appropriate phase conditions between the coupled modes.

Directional couplers are useful in numerous optical components. For example, optical switches in Mach-Zehnder Interferometer (MZI) configuration and ring-resonator-based optical filters utilize DCs. The wavelength dependence of the DC allows for spectral MUX/DEMUX of different wavelength signals. DCs are also used in grating assisted devices for efficient WG couplers [21]. In fact, one can find optical DCs since integrated optics was born, and they have been demonstrated and tested in every available material system—$LiNbO_3$, silica on Si, polymers, glass, GaAs, InP, and in the last decade also on SOI. In general, DCs using silica buried (channel) WGs provide the lowest loss (0.05–0.1 dB/coupler), but their length is relatively large (on the order of millimeters) due to the low core-cladding index contrast (~10^{-2}—a relative index contrast of ~0.75%), which requires bend radii of more than 5 mm for low-loss behavior. In addition, their polarization dependent loss (PDL) is very low and negligible.

27.2.3.2. Multimode Interference Coupler (MMI)

The multimode interference (MMI) coupler [22] is a photonic device that employs the well-known Talbot effect leading to self-imaging [23, 24] in multimode waveguides and has become a key component for photonic integrated circuits. The central structure of a MMI coupler is the multimode waveguide that can support a large number of guided modes (typically more than three modes). Figure 27.6 shows schematically a NxM MMI coupler with N single-mode input

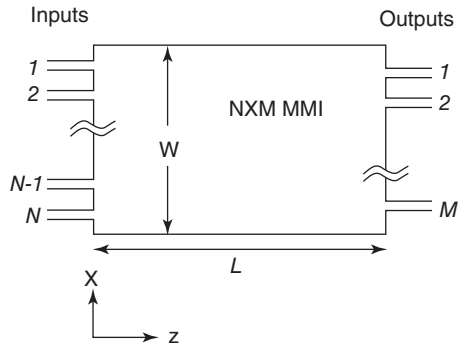


Figure 27.6 Schematic of a NxM MMI coupler with N input waveguides and M output waveguides. The multimode waveguide has a width of *W* and length of *L* and supports a large number of modes.

waveguides and M single-mode output waveguides connected to the multimode waveguide having a width of W and a length of L. The MMI coupler can be used for directional coupling, splitting, and combining. The main benefit of an MMI device compared to conventional directional coupler (see above in 27.2.3.1) is its better tolerance to fabrication insensitivities due to no need for a gap between the waveguides that may put severe requirements on the fabrication, especially in very small gaps (fraction of a micron). In addition, MMIs are polarization independent which may significantly improve the performance of a polarization independent device. On the other hand, MMIs may show higher excess loss, but with the correct design, one can overcome this and obtain reasonable loss.

Silicon waveguide-based MMI couplers can be used in many optical components such as the optical modulator, optical switch, and ring resonator. For these applications, the MMI device works as a 50–50% splitter/combiner, just like a 3-dB directional coupler. The first SOI waveguide based 1×2 and 2×2 MMI couplers were reported by Fischer, Zinke, and Petermann in 1995 [25]. These devices were used to form 1×2 and 2×2 optical switches. Silicon waveguide 1×4 and 1×8 MMI couplers were also reported [26]. Good power splitting and low on-chip loss were obtained. The MMI coupler has also been used in an arrayed-waveguide grating to broaden the spectral response [27] and variable optical attenuator [28].

27.2.3.3. Y-Junction

The Y-junction is a simple 3-ports device in which one single-mode (SM) WG branches into two SM WGs as shown in Fig. 27.7. A tapered waveguide region bridges the single WG and the two WGs. In order to have small excess loss, the tapered waveguide must be able to support more than one guided modes—typically, both the symmetric and the antisymmetric eigenmodes.

The Y-junction can operate as either a splitter (propagating from the single WG to the two WGs) or a combiner (propagating from the two WGs to the single WG). When the junction is designed symmetrically (as in Fig. 27.7), it performs as a 3-dB splitter. When light mode is injected through the one WG input, due to the geometrical symmetry of the structure and the adiabatic design of the taper,

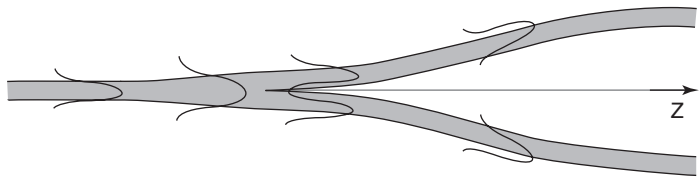


Figure 27.7 Top view of an optical Y-junction.

only symmetric mode is excited in the tapered region. Hence an equal power split into the two branches occurs with little excess loss.

Y-junction splitters can also be designed to realize weighted splitting or even optical taps. For instance, one can modify the width of the two output WGs to differ from each other or to change their angle with respect to the z-axis to obtain broadband and polarization independent asymmetric splitters.

Y-junctions can find numerous applications. Working as a 3-dB splitter/combiner, the Y-junction is frequently used to construct MZIs. Because the MZI requires very accurate 50–50% splitting with wavelength insensitivity and polarization independence, the Y-junction-based MZI typically has a robust performance. The Y-junction can also be used for constructing the $1 \times N$ splitter. A multiplication of concatenated 3-dB splitters (as a binary tree) of N levels can split a single input into 2^N equal outputs. Lastly, the Y-junction is useful for digital optical switch (DOS). Optical switches employing interferometric mechanism possess strong wavelength and polarization dependence, which requires specific and quite accurate voltages per each element due to unavoidable fluctuations in fabrication. The DOS switch [29–31], however, has a step-like nonperiodic switch response and is much less wavelength and polarization dependent. This digital (binary) characteristic can provide a unique and very attractive and tolerant element (wide window of operational parameters and process). As a very fundamental building block, a silicon waveguide-based Y-junction was designed and fabricated in the early stage of silicon photonics research (for example, see Refs. [32–36]).

27.2.4. Mach-Zehnder, Gratings, and Ring Resonators

27.2.4.1. Mach-Zehnder Interferometer

The Mach-Zehnder interferometer (MZI) is a most versatile and widely used device in which a light beam is split into two parts and then each part undergoes a different physical path that may create a phase difference, and finally the two parts are combined to provide the end result where the phase difference is converted into a change in output light intensity. The MZI is extensively utilized in numerous forms and materials and for countless applications. In integrated optics there are mainly three basic MZI configurations, as illustrated schematically in Fig. 27.8.

1×1 MZI [Fig. 27.8(a)] makes use of a 1×2 splitter and a 2×1 combiner usually based on Y-junctions or MMIs. 2×2 MZI [Fig. 27.8(b)] consists of two 2×2 3-dB couplers based on DCs or MMIs. The third type of MZI [Fig. 27.8(c)] can be defined as 1×2 or 2×1 MZIs and is a combination of the first two configurations.

The operation principle of the 1×1 MZI is relatively simple. The input light is split into two equal parts; each propagates in its own WG arm. The two WG

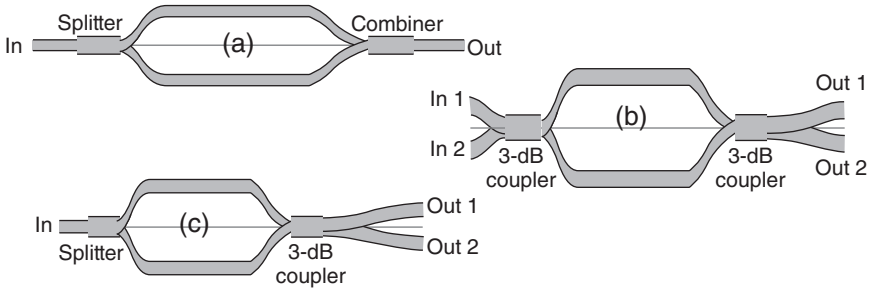


Figure 27.8 Three main MZI configurations (a) 1×1 MZI, (b) 2×2 MZI, and (c) 1×2 MZI.

arms may have the same (symmetric) or different (asymmetric) optical path length (OPL) due to possible effective refractive index change or waveguide length difference. If light in the two arms acquires no phase difference between them, they arrive at the combiner as a symmetrical mode; that is, they recombine in-phase (constructive interference), and the lightwave fully emerges from the output SM WG—*on state*. When the two beams acquire a phase difference of π, only an antisymmetrical eigenmode is produced, and it is totally lost in the combiner output SM WG—*off state*. In the case where the two WG arms differ in their propagation losses, a nonequal arm loss ingredient would degrade the extinction ratio (ER) defined as $10 \log_{10}(P_{on}/P_{off})$. When the two arms of the MZI have a waveguide length difference (unbalanced MZI), ΔL, the optical phase difference is $\Delta\phi = 2\pi n_{eff}\Delta L/\lambda$, where n_{eff} is the effective index. Thus, for the unbalanced MZI configuration, one expects a wavelength dependence of the output intensity.

In the case of 2×2 MZI, after the first 3-dB coupler, light from one of the input ports is equally split into two beams, and the beam in the cross port lags behind the beam in the bar port by a phase of $\pi/2$. When the arms have no phase shift, the two equal power beams arrive at the second coupler with $\pi/2$ phase difference. These are exactly the same conditions the light would "see" in the middle of a full coupler (of length L_c); thus, the second 3-dB coupler can be considered as a direct continuation of the first one, and the two work together as a single coupler with full coupling length. All the light therefore propagates through the MZI out of its cross output (cross state). If the difference in OPL between the two arms is of $\lambda/2n_{eff}$, that is, a π phase shift, then the phase conditions between the two arms at the input to the second 3-dB coupler are exchanged in regard to the previous symmetrical case, and the second coupler acts in reverse to the first one. Therefore, the lightwave goes back to its original WG—the MZI is in bar state. It is clear then that the cross state requires that the combination of the two couplers (continuation in coupling direction) will provide an exact full coupling ($X + Y = 1$, where X and Y are the two power coupling ratios in the first and the second couplers), whereas the bar state (reverse operation) requires only that the two

couplers will be identical in coupling ratios ($Y - X = 0 \Rightarrow X = Y$). Therefore, in order to allow both states, one requires that the coupling ratio will be identical but with a specific ratio of 3-dB couplers ($X = Y$ and $X + Y = 1 \Rightarrow X = Y = 1/2$). The identical requirement can be easily obtained with very good process tolerance since the two couplers are in relatively close proximity on the wafer. However, the exact 3-dB coupler requirement (or any other specific value) is much harder to achieve with standard process capabilities. As a result, the bar state can be achieved more easily than the cross state and its cross talk (CT) is usually much better than in cross state (approx. –30 dB vs. –15 dB). Wavelength dependence of the 2×2 MZI comes from the 3-dB coupler wavelength dependence and also from the OPL difference, which depends on λ. We expect a wider spectral response at bar state (requiring only identical couplers) than at cross state (requiring exact 3-dB couplers). The cross-state wavelength insensitivity can be improved by using wavelength insensitive couplers, but these are usually more complicated, increasing loss and length compared to the conventional 3-dB couplers, and should be used only when necessary.

The operational principle of the third type—combined configuration (Fig. 27.8(c)—splitter and 3-dB coupler) is as follows. First, light is equally split by the splitter with no phase difference (in phase). In this case, the input to the 3-dB coupler is only a symmetrical eigenmode; hence, the total output is also symmetrical. Therefore, the 1×2 MZI with no OPL difference in arms acts as a 3-dB coupler in its passive state. When the two arms have a phase difference of $\pi/2$, all the light will emerge from one output WG, and when the phase difference is $-\pi/2$, all the light will emerge from the second output WG, both similarly to the cross state of the 2×2 MZI. The main difference between the 1×2 MZI and the 1×1 or 2×2 MZIs is in its smaller required maximum power values (half) due to the half required phase shift ($\pi/2$ vs. π). Yet the average power in time is similar since the 1×2 requires a constant activation and the others consume power only in about half of their life. In addition, the 1×2 configuration lacks the process immunity of the bar state in the 2×2 MZI.

A very important feature can be achieved with an MZI; in addition to the full cross or bar states (binary), one can tune the MZI to get any intermediate state to allow analog output conditions. This can be obtained by inducing intermediate-phase conditions, enabling various applications like weighted multicasting and weighted equalization at the output ports of an optical circuit [37].

As expected, the MZI is also one of the most investigated devices in silicon photonics. The most widespread and well-developed element is the Si modulator based on 1×1 MZI but with various different modulation mechanisms employed on one or both of the WG arms. For a modulator, the thermo-optic (TO) effect is too slow (~KHz), whereas the free carrier plasma dispersion effect can be much faster (MHz and GHz) and can be realized in numerous ways. The first is a forward-bias PIN diode [18, 38, 39] in which an operation of 10–20 MHz was

achieved. The GHz boundary was first passed by Intel [40] and later on was extended to 10 GHz [1]. Both were based on forward-biased MOS capacitors operating in the very fast accumulation state. Recently, a high-speed modulator based on the reverse-biased PN diode was designed and fabricated [41]. Using such a modulator, data transmission with bit rate of 40 Gb/s was demonstrated.

Other 1×1 MZIs were also suggested based on the TO effect; see, for example, in [34] and [42] on Si-wire WGs. The spectral filtering capability of MZIs was also demonstrated on silicon WG technology. In [43] Jalali et al. demonstrated a wavelength splitter based on Si 2×2 DC-MZI that splits two wavelengths within the C-band with FSR of 8 nm and CT of −18 dB. Liu et al. [44] have demonstrated a MZ wavelength combiner based on a 2×2 MMI-MZI device for Raman amplification ($\lambda_1 = 1440$ nm, $\lambda_2 = 1556$ nm) with wavelength selectivity of 20 dB, insertion loss (includes fiber-WG couplings) of 9 dB, and PDL of 0.8 dB. A 4×4 switching matrix in SOI based on a fabric of 5 TO 2×2–MMI–MZI elements was considered by Wang et al. [45] and by Li [46], demonstrating at $\lambda = 1550$ nm a CT of −12 dB to −20 dB with on-chip path loss of 6.6–10 dB and switching time of less than 30 μs. The power consumption per switch element was 330 mW.

27.2.4.2. Waveguide Bragg Gratings

Bragg gratings [47] are spectral filters that typically reflect light in a narrow band of wavelengths and allow light transmission in all other wavelengths. They are based on the principle of Bragg reflection and are formed by creating a periodic structural corrugation or material refractive index modulation along the length of a waveguide. Figure 27.9 shows the schematics of two waveguide gratings. One has shallow corrugations etched into the top surface of the rib waveguide, and the other has alternating regions of higher and lower refractive indices built into the rib waveguide.

Each unit or period of the grating, due to its structural or index discontinuity, acts as a weak reflector where light would experience both reflection and transmission. If the length of each period, Λ, is such that all the partial reflections add up in phase, which occurs when the round trip of the light between two reflections

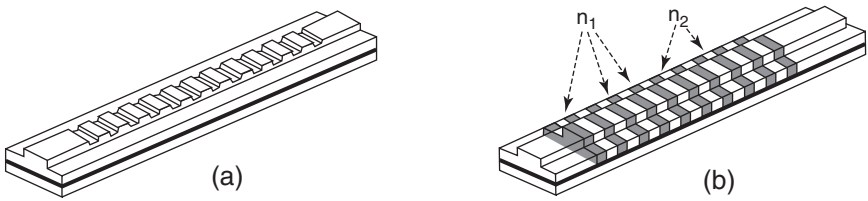


Figure 27.9 Schematics of two waveguide grating designs: (a) is based on surface corrugations and (b) is based on refractive index modulation.

is an integer multiple of the wavelength, the Bragg condition is satisfied and the total reflection can sum up to be nearly 100%. The Bragg phase-matching condition can therefore be expressed as

$$2\beta_B\Lambda = 2m\pi \tag{27.2}$$

where m is an integer denoting the grating order and β_B is the propagation constant of the optical wave meeting the Bragg condition, defined as

$$\beta_B = \frac{2\pi n_{eff}}{\lambda_B} \tag{27.3}$$

where n_{eff} is the effective refractive index of the mode guided in the grating structure and λ_B is the Bragg wavelength in free space. By substituting Eq. (27.3) into Eq. (27.2), the Bragg condition simplifies to

$$\Lambda = m\lambda_B/2n_{eff} \tag{27.4}$$

A first-order ($m = 1$) Bragg grating is usually analyzed using the coupled mode theory that describes the coupling between the forward and backward propagating waves at a given wavelength. For a uniform grating in which the refractive index varies along the grating longitudinal axis z as $n(z) = n_{ave} + \Delta n \cos(2\pi z/\Lambda)$, the grating amplitude reflection coefficient is given by

$$r_g = \frac{i\zeta \sin(qL)}{q\cos(qL) - i\Delta\beta \sin(qL)} \tag{27.5}$$

where $\Delta\beta = \beta - \beta_B$ is the propagation constant detuning away from the Bragg wavelength, $\zeta = 2\pi\Delta n/\lambda$ is the coupling coefficient that is related to the index modulation depth, $q = \sqrt{\Delta\beta^2 - \zeta^2}$, and L is the grating length. The grating power reflectivity is given by $|r_g|^2$. Note that the waveguide loss is not included in the above analysis. Equation (27.5) suggests that the grating reflection spectrum strongly depends on the grating coupling coefficient or the refractive index modulation depth and grating length. The grating spectral bandwidth, centered on the Bragg wavelength, as well as the grating reflectivity, can therefore be tailored to meet the intended application requirement. To achieve a narrow bandwidth, the grating coupling strength needs to be weak so that the reflection by each grating period is small. This can be designed by reducing the cross section of the structural corrugation, reducing the periodic material refractive index contrast, or reducing the overlap of the optical mode with the grating. For such weak gratings, bandwidth is also inversely proportional to the grating length, so the longer the grating, the narrower the bandwidth. Some notable applications of waveguide Bragg gratings are optical filtering, optical add-drop multiplexing (OADM), and the enabling of narrow linewidth semiconductor lasers.

In recent years, researchers in the field of Si photonics have proposed and demonstrated numerous waveguide-based Bragg gratings in Si. Because most Si

waveguides have effective refractive indexes greater than 3, the first-order Bragg grating period must be <0.25 μm to achieve a Bragg wavelength of 1.55 μm. This small feature size makes device processing very challenging because the lithographic resolution needed is <0.13 μm. While 0.13 μm lithography is common in today's IC Fabs, it is not readily available to most optics researchers. Many groups have therefore turned to E-beam lithography to pattern their gratings [48–50]. A group has also recently experimented with focus ion beam milling to define their devices [51]. All of these works formed their gratings by etching structural corrugations either into the top surface or sidewalls of the waveguides.

27.2.4.3. Ring Resonators

Another resonator design that has gained widespread interest is the ring resonator (RR), which is functionally the same as the F-P resonator. The difference is that it consists of a waveguide in a closed loop, commonly in the shape of a ring or racetrack. Also, the RR provides traveling wave operation, as opposed to the standing wave operation characteristic of Fabry-Perot resonators. As Fig. 27.10 illustrates, light can be coupled into the ring via evanescent field coupling by placing the input waveguide, also known as the input bus, within close proximity of the ring. Like the F-P, the ring behaves as an interferometer and will be resonant for light whose phase change after each full trip around the ring is an integer multiple of 2π, which is of course the condition for the light in the ring to be in-phase with the incoming light and constructive interference. Light that does not meet this resonant condition is transmitted through the bus waveguide. The expression for the resonant wavelengths of the ring is very similar to that of the F-P and is given by

$$\lambda_r = \frac{2\pi R n_{eff}}{m} \tag{27.6}$$

where R is the ring radius with circular waveguide ring, and m is an integer.

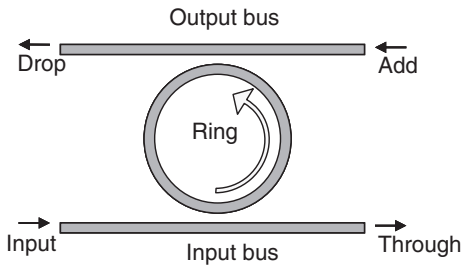


Figure 27.10 Top-down schematic view of a ring resonator.

When the RR is coupled to a single waveguide, the transmission response of the resonant wavelengths through the bus strongly depends on the optical loss of the ring and the coupling efficiency between the ring and bus waveguide. The device can act as a phase filter where all wavelengths are transmitted and the resonant wavelengths, having also traversed the ring, acquire a phase change. Or, the device can exhibit notch filter behavior where all the light on-resonance is coupled to the ring, giving high extinction in the transmission spectrum. To capture or separate the resonant wavelengths from the rest, an additional waveguide, an output bus, can be placed on the opposite side of the ring as shown in Fig. 27.10. If the optical loss in the ring is negligible, all the on-resonance light that is coupled into the ring can be coupled into the output bus toward the drop port. This design configuration is actually even more versatile; it can have all the functionality of an optical add-drop multiplexer (OADM) because wavelengths can also be added via the "add" port, get coupled into the RR, and exit the "through" port via the input bus waveguide.

Key performance parameters of the RR include the FSR, the ER, and the finesse. The expression for the FSR of a RR is given by

$$\Delta\lambda = \frac{\lambda_r^2}{2\pi R n_g} \tag{27.7}$$

To calculate the finesse or Q factor, one must of course first define the 3-dB bandwidth. Assuming the coupling between the ring and the two bus waveguides are weak and can both be represented by the coupling coefficient k, the 3-dB bandwidth can be approximated as $\delta\lambda \approx \kappa^2\lambda_r^2/(2\pi^2 R n_g)$. When the RR is on resonance, light coupled into the ring constructively interferes with the input light. As a result, optical intensity in the ring can build up and be significantly higher than that in the bus waveguide. This field enhancement is an important property of RRs and can be measured by its finesse or Q factor. As light makes multiple round trips in the ring, it will also experience loss—transmission loss of the ring waveguide and loss due to coupling to the bus waveguides. If N is the number of round trips required to reduce the optical energy to $1/e$ of its initial value, finesse is given by

$$F = 2\pi N \tag{27.8}$$

and the Q—quality factor is given by

$$Q = \omega_r T N \tag{27.9}$$

where ω_r is the resonant frequency and T is the time it takes to make one trip around the ring. It is therefore easy to see that to achieve high field enhancement, high finesse or Q, the transmission loss of the ring must be minimized and the coupling efficiency must be optimized.

SOI-based RRs have been fabricated and successfully demonstrated by numerous research teams. Many are based on very compact, single-mode strip waveguides whose core crosssections are approximately 200 nm x 500 nm. In SOI systems, where optical confinement is strong due to the large refractive index contrast between the Si core and SiO_2 cladding, such small core dimensions ensure that very compact rings with bending radii on the order of a few micrometers can be realized without significant bend loss [52–59].

An optical buffer or an optical delay line is a crucial component both for computer optical interconnects and for Datacom optical routers. This is an essential optical memory that can prevent signal congestion and eliminates the need for high-cost O-E-O conversions. Recently, F. Xia and co-workers from IBM [60] have demonstrated an interesting candidate for such optical buffers. It is based on cascading silicon ring resonators delay lines. They produce a group delay of 10 bits at a bit rate of 20 Gbps in an on chip area smaller than 0.1 mm^2.

27.3. ACTIVE SILICON PHOTONIC DEVICES

In this section, we describe some selected active elements and devices such as modulators, photodetectors, amplifiers, and lasers. In active devices we use the wide-sense definition of elements that employ refractive index tuning in addition to the strict version of elements with gain. If one can refer to the history of silicon photonics as two phases, then the second one (last ~6 years) is more characterized by breakthroughs in active silicon devices, in addition to evolutionary improvements in passive device performance. These active devices were the missing link in siliconizing photonics, and the jewel in the crown is with no doubt the silicon laser.

27.3.1. Refractive Index Control

Control mechanisms of the refractive index (RI) are key enablers for any guided wave technology. The intensity and speed of the RI modification effect and its simplicity are major parameters in the success/failure of the technology. As expected, any change in RI is coupled to a modification in the absorption (Kramers-Kronig relations) and vice versa. Silicon as a centro-symmetric crystal does not exhibit the attractive electro-optic Pockels effect. In addition, its quadratic (Kerr) electro-optic effect and its electroabsorption (Franz-Keldysh) effect are too small. As a result, until recently (~2004) no fast RI tuning effect was available, and this was a major obstacle for silicon photonics. This obstacle was overcome, as will be described in Section 27.3.2. The main two strong RI tuning effects are the thermo-optic (TO) effect, which is naturally slow (KHz), and the free carrier plasma-dispersion effect that can be employed both slowly (MHz) and very fast (GHz).

The TO effect in Si WGs is based on the very high polarizability of silicon and bandgap dependence on temperature, which creates a TO coefficient of dn/dT

$\cong 1.86x10^{-4}$ (1/°*K*) that is more than 15 times higher than in SiO_2 or SiON materials. In addition, thermal conductivity is much higher in silicon than in other materials—about 100 times higher than in SiO_2, SiON, or polymers. Both the thermal conductivity and the TO coefficient provide a much faster and more efficient TO effect in silicon than in other materials.

The free carrier plasma-dispersion effect is the most utilized and the one that allows the Gb/s regime. It is based on free carriers (holes and electrons) injection/depletion changing both absorption and RI, as the imaginary and the real parts of the complex RI, and coupled by the Kramers = Kronig dispersion relations [61, 62].

27.3.2. Modulators

Replacing the metal interconnect with optical lines will decrease the delay and will allow higher bandwidth that is required due to transistor speed improvements, and potentially will lower the power consumption. One of the main elements to allow optical interconnect is the optical modulator, converting efficiently an electrical signal into an optical version at high bit rates (10–40 Gbps). Today's commercially available high-speed optical modulators at >10 Gb/s are based on electro-optic materials such as lithium niobate [63] and III–V semiconductors [64]. These devices have demonstrated modulation capability as high as 40 Gb/s. Achieving fast modulation in silicon was a challenging task.

27.3.2.1. Si MOS Modulator design [40, 1]

During 2004–2005, a metal-oxide-semiconductor (MOS) capacitor-based Si modulator was proposed and experimentally demonstrated a modulation bandwidth of 2.5 GHz and 10 GHz at optical wavelengths of around 1.55 μm. Its operation is based on the free-carrier plasma dispersion effect wherein the phase-shifting elements of an MZI are MOS capacitors embedded in Si rib waveguides. An applied voltage induces an accumulation of charges near the gate dielectric of the capacitor, which, in turn, modifies the refractive index of the waveguide and ultimately the optical phase of light passing through it.

Figure 27.11 shows a cross section of the silicon waveguide MOS capacitor-based phase shifter. It comprises an optimized $1.7 \times 10^{16}/cm^3$ n-type doped crystalline silicon slab and $5 \times 10^{16}/cm^3$ p-type doped polysilicon rib with 120 Å gate oxide sandwiched between them.

Figure 27.12 is a cross-sectional scanning electron microscope (SEM) image of this new MOS capacitor phase shifter. It has a smaller cross section of $1.6 \times 1.6\,\mu m^2$, which comprises a 1.0-μm n-type doped crystalline Si on the bottom and a 0.55-μm p-type doped crystalline Si on the top, with a 10.5-nm dielectric sandwiched between them.

The transition to crystalline silicon was done by introducing an epitaxial lateral overgrowth (ELO) process which enabled forming the crystalline Si on top of the

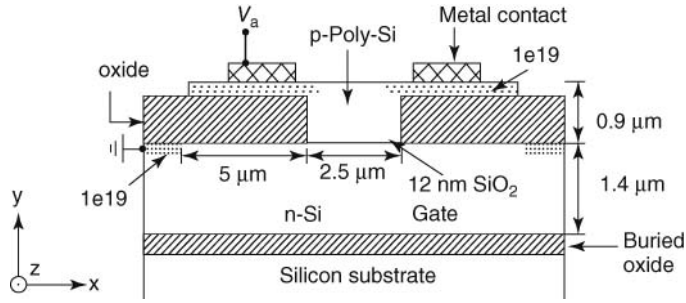


Figure 27.11 Cross-sectional view of the MOS capacitor phase shifter. The optical mode propagates in the z direction.

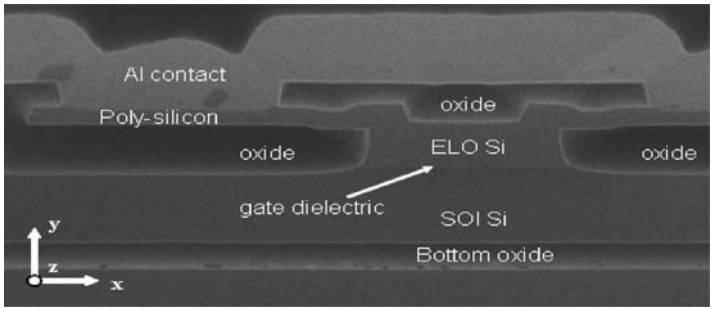


Figure 27.12 SEM cross-section of the new phase shifter.

gate dielectric. With low-loss material, we increased doping concentration (keeping the transition loss < 10 dB) to $2 \times 10^{17}/cm^3$ for the Si slab and $1 \times 10^{18}/cm^3$ for the ELO-Si rib dimensions. By changing the dimensions and decreasing the distance between the metal contacts and the waveguide rib by >40% we managed to achieve a bandwidth of 10 GHz and transition loss of 9 dB/L_π.

In the new design [1], all dimensions are smaller, and as a result, the optical mode is more tightly confined and interacts more strongly with the voltage-induced charges. The phase efficiency, $V_\pi L$ product, of the new device is therefore improved, 3.3 V-cm compared to 7.8 V-cm. Despite the new device's significantly higher dopant concentrations, its transmission loss remained relatively constant at 9 dB/L. This is the combined result of the improvement in phase efficiency and the use of the lower loss ELO-Si.

27.3.2.2. PN Carrier Depletion Silicon 40 Gb/s Modulator [41]

Next we present an ultra-fast silicon modulator based on the carrier depletion effect in a reverse-biased pn junction. To reduce the RC constant limitation, a

traveling-wave electrode based on coupled coplanar waveguide and microstrip is designed. In particular, we demonstrate the high-speed modulation with data transmission of 40 Gb/s.

Figure 27.13(a) shows schematically the MZI modulator with a reverse-biased pn diode embedded in each of the two arms. To obtain better phase-modulation efficiency, we designed and fabricated a submicrometer-size waveguide. Silicon rib waveguide width is ~0.6 μm, rib height is ~0.5 μm, and etch depth is ~0.22 μm. Both modeling and experiment confirm that the waveguide is a single-mode device for wavelengths around 1.55 μm. The waveguide splitter is a 1×2 multimode interference (MMI) coupler. The key active component of the silicon modulator is the reverse-biased pn junction phase shifters embedded in the MZI arms.

Figure 27.13(b) shows a schematic of the cross-sectional view of the phase shifter. It comprises a p-type doped crystalline silicon rib waveguide having a rib width of ~0.6 μm and a rib height of ~0.5 μm with an n-type doped silicon cap layer (~1.8 μm wide). This thin (~0.1 μm thick) cap layer is formed using a nonselective epitaxial silicon growth process and is used for pn junction formation and electrical contact. The p-doping concentration is ~1.5×10^{17} cm^{-3}, and the n-doping concentration varies from ~3×10^{18} cm^{-3} near the top of the cap layer to ~1.5×10^{17} cm^{-3} at the pn junction. The process is designed to target the pn junction at approximately 0.4 μm above the buried oxide to enable optimal

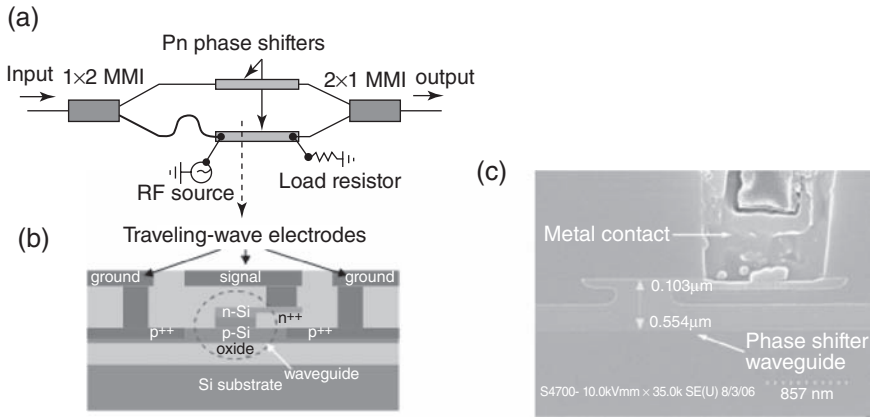


Figure 27.13 (a) Top view of an asymmetric Mach-Zehnder interferometer silicon modulator containing two pn junction based phase shifters. The waveguide splitter is a 1×2 multimode interference (MMI) coupler. The RF signal is coupled to the traveling wave electrode from the optical input side, and termination load is added to the output side. (b) Cross-sectional view of a pn junction waveguide phase shifter in Silicon-On-Insulator. The coplanar waveguide electrode has a signal metal width of ~6 μm and a signal-ground metal separation of ~3 μm. The metal thickness is ~1.5 μm. The high-frequency characteristic impedance of the traveling wave electrode is ~20 Ω. (c) Scanning electron microscope (SEM) image of a pn diode phase shifter waveguide.

modal overlap with the depletion region. As the n-doping concentration is much higher than the p-doping concentration, carrier depletion under reverse bias occurs mainly in the p-type doped region. This leads to better phase modulation efficiency because the hole density change results in a larger refractive index change as compared to the electron density change according to Kramers-Kronig analysis of optical absorption spectrum [62]. The cross-sectional scanning electron microscope (SEM) image of a fabricated pn diode phase shifter waveguide is shown in Fig. 27.13(c).

For a reverse-biased pn junction, the depletion width depends on the bias voltage and doping concentrations. For the asymmetrically doped pn junction in this work (the n-doping concentration is much higher than the p-doping concentration), the depletion width (W_D) can be approximated by $W_D = (2\varepsilon_0\varepsilon_r(V_{Bi} + V_{app})/eN_A)^{1/2}$ where ε_r is the low-frequency relative permittivity of silicon, N_A is the acceptor concentration, V_{Bi} is the built-in voltage, and V_{app} is the applied voltage. Note the nonlinear dependence of W_D on the bias. Changing the depletion width of a pn junction is equivalent to changing the free carrier density. Thus, by changing the bias voltage, one can achieve refractive index modulation through the free carrier plasma dispersion effect.

27.3.3. NIR Si-Ge Photodetectors

At the interface between every optical link and electronic central process unit, data storage, or OEO node, one will find an optical electrical receiver that contains an electrical IC (TIA, pre amp) usually from Si CMOS technology and a III–V-based photodetector. In the Telecom region for the Metro and long haul, the data transfers on the 1550-nm and 1310-nm optical carrier channel and the vast majority of the photo detectors used in the market is based on InGaAs [65]. For shorter ranges of optical link in the Enterprise region for high-speed router, server and storage, FTTH, and more, the market volume is much larger and the focus is on low-cost optical components. Therefore the optical carriers are mostly 850 nm and 1310 nm. For the 850 nm, the most common photodetectors are made of GaAs on InP or on GaAs substrates. The Si-based receiver was chosen many years ago as the right candidate for significant reduction in cost of the receiver. However, most communication-grade semiconductor lasers are operating in the near infrared (NIR) wavelength (usually 0.850, 1.31 and 1.55 μm), a region where silicon is a poor detector. In order to improve the performance of silicon-based detectors, the most common approach is to introduce germanium to reduce the band gap and extend the maximum detectable wavelength. The effect on the absorption coefficient and penetration depth, defined as distance that light travels before the intensity falls to 36% (l/e), is clearly shown in Fig. 27.14.

Two critical benchmarks for a photodetector are directly related to the absorption coefficient or penetration depth of the light: responsivity and bandwidth

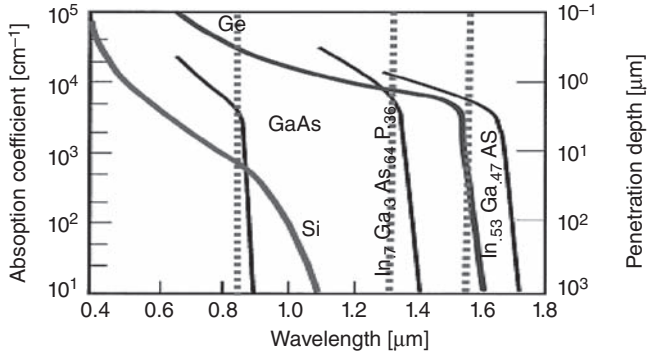


Figure 27.14 Absorption coefficient and penetration depth of various bulk materials as a function of wavelength. The dotted lines mark the important wavelengths for telecommunications of 0.85, 1.31, and 1.55 μm.

(BW). The responsivity is the ratio of collected current to the optical power incident on the detector. Responsivity of commercial GaAs photodetectors is typically close to 0.55 A/W (at 850 nm); the responsivity should increase as the absorption coefficient increases. The bandwidth of a photodetector can be limited by the transit time required for the carriers to travel to the contacts or the RC time constant of the circuit. The RC time constant (t_{RC} = RC) is proportional to 1/d (when d is the intrinsic layer thickness of PIN photodetector) because the capacitance part, C = εA/d (ε is the dielectric constant of the intrinsic layer, and A is the diode area). The transit time, $t_{tr} = d/V_d$ (V_d is the drift velocity of the free carrier), is proportional to d in case of saturated velocity (high field) and to d^2 in case of low field. Therefore, we have optimum point for the temporal response, see Fig. 13, while the responsivity increases with d till saturation. If the penetration depth can be kept below 2.5 μm (and above 1 μm), the transit time alone can support a bandwidth of 10 Gb/s for 50-μm PD diameter. We approximately combine the two contributors to the total BW to be

$$BW = [(1/BW_{tr})^2 + (1/BW_{RC})^2]^{-1/2} \quad (27.10)$$

For a real device, the data will be influenced by series resistance and parasitic capacitance. Our Ge on Si PIN photodetectors with 1.5-μm intrinsic Ge show measured 3-dB bandwidth of 9 GHz.

27.3.4. Silicon Amplifiers and Lasers

27.3.4.1. Introduction

Although significant progress has occurred in recent decades, the electrically pumped silicon laser has been and remains the major obstacle and the holy grail

that may revolutionize the fields of optoelectronics, communication, and computing. This is a most challenging issue because silicon is an indirect bandgap semiconductor, requiring a phonon for electron-hole recombination. Hence, the radiative emission lifetime is much longer than nonradiative processes (impurities traps and Auger recombination). As a result, the internal quantum efficiency, $\eta_i = r_{rad}/(r_{rad} + r_{nr}) = \tau_{nr}/(\tau_{rad} + \tau_{nr})$ (where r denotes rate (1/s), τ stands for lifetime (sec), and the subscripts *rad* and *nr* indicate radiative and nonradiative processes, respectively) is ~10^{-5} in silicon. In addition, the free carrier absorption (described in Subsection 27.3.1) in silicon is much higher than the low amplification that theoretically could be obtained, resulting in net optical loss. This gloomy status is totally opposed to direct semiconductor materials like GaAs and InP, which control most of the semiconductor laser market. Three main methods have been proposed to obtain the first silicon laser:

1. *Quantum confinement* of charge carriers based on low-dimensional structures that reduce significantly the rate of nonradiative recombination (related to the uncertainty principle—increased momentum uncertainty of charge carriers due to their localization). A popular technique is to employ Si Nanocrystals (Si-nc) fabricated through high-temperature annealing (~1200°C for several minutes) of substoichiometric silica films (SiO_x, $x < 2$). Si-nc has been tested and has provided [66, 67] promising photo/electroluminescence. However, some results need further explanation, results are not always reproducible, and the operating wavelength range of Si-nc (800–1000 nm) is outside of the long-haul optical communication window (around 1310 nm and 1550 nm—O, C, and L bands).

   Another quantum confinement technique is the suggested porous silicon of easily attained isolated silicon pillars via an exposure to HF acid. The pillars' width dimension is a few nanometers allowing quantum confinement [68]. LED operation with 1% was achieved at a quantum efficiency bias of 5 V Voltage. However, its reliability performance should be much improved.
2. *Erbium doping* Following the success of erbium doped fiber amplifier (EDFA) in optical fiber communication, silicon was implanted with Er^{3+} ions and annealed. The Er ions were exited optically [69, 70] similar to EDFAs, and also electrically based on transfer of charge carrier recombination energy to the Er. Although sufficient Er luminescence is obtained at low temperatures ($T < 100$ K), a severe signal quenching is observed at room temperature. This degradation is explained by a reverse mechanism whereby Er ions nonradiatively back transfer their energy to defect levels in the bandgap through free electrons. The quenching effect can be alleviated by co-doping (ion implantation) oxygen ions together with the Er.

A combination of nanocrystals with erbium was also tested [71] and provided better luminescence efficiency than the Er + O implanted bulk.

3. *Stimulated Raman scattering* (SRS) is a third technique to achieve the silicon laser based on the Raman effect and is described in the next subsection. Although this method succeeded in producing the first silicon laser, it is optically pumped by another laser and not electrically pumped.

27.3.4.2. Raman Amplifiers and Lasers

One of the more recent and important branches in silicon photonics is the study of various nonlinear optical effects in silicon to create active photonics devices. Due to silicon's high-index contrast to its oxide, light can be confined in silicon waveguides more tightly than in glass fiber or silica-based waveguides. The optical modal area in silicon waveguides can be more than 10^3 times smaller than in a single-mode fiber. Moreover, many optical nonlinear effects are much stronger in silicon than in glass fibers; for instance, the Raman gain coefficient in silicon is 3–4 orders of magnitude greater than in an ordinary glass fiber. Many of the concepts and applications already developed today that are based on nonlinear optical effects in silica fibers can be adapted for silicon waveguides to produce compact, chip-scale active photonics devices for various applications, including optical amplifiers, lasers, and wavelength converters, as well as ultra-fast switching, pulse shaping, and devices based on slow-light generation.

Although silicon has a much higher Raman gain coefficient and silicon waveguides provide stronger light confinement when compared to silica fiber, the optical loss in silicon is significantly higher than in fiber, especially the nonlinear loss due to two-photon absorption (TPA) induced free carrier absorption (FCA) [72, 73]. Net optical gain was achieved for the first time in silicon waveguides using a pulsed pump configuration [74], and with narrower pump pulses, on-off peak Raman gains as high as 20 dB have been reported [75]. CW amplification and lasing are significantly more challenging due to the competing nonlinear optical loss mechanism, as TPA-induced FCA causes the optical loss to increase with pump power. By introducing a reverse-biased p-i-n diode structure embedded in a silicon waveguide, the nonlinear absorption can be efficiently reduced and continuous-wave net Raman gain has been demonstrated [76]. Figure 27.15 is a simple schematic of an all-optical amplifier; a pump source is directed into the silicon waveguide and is combined with the incoming datastream passing through the silicon waveguide. By using this configuration, on-chip amplification of nearly ~3 db of a 10 Gb/s datastream and on-off gain of 2.3 dB of a 40 Gb/s data stream has been recently demonstrated [77]. As a nonlinear effect, optical pump-induced stimulated Raman scattering (SRS) provides a means to generate optical gain and lasing in silicon.

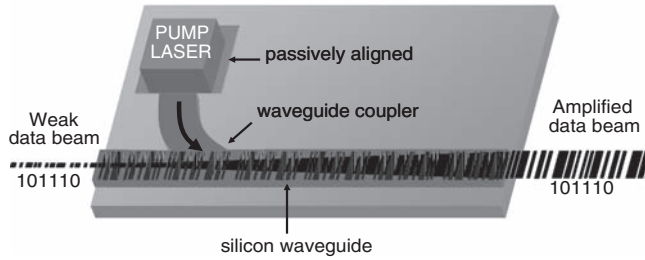


Figure 27.15 Simplified schematic of a silicon-based on-chip optical amplifier.

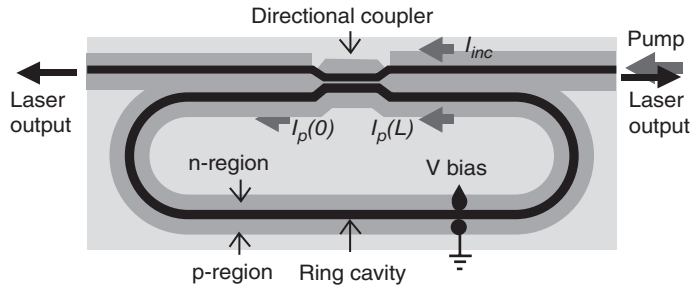


Figure 27.16 Layout of the silicon ring laser cavity with a p-i-n structure along the waveguides.

Linear Cavity Laser

The CW silicon Raman laser optical cavity was formed by coating both facets of the low-loss silicon waveguide with multilayer, dielectric films. The front facet coating is dichroic, having a reflectivity of ~71% for the Raman/Stokes wavelength of 1686 nm, and ~24% for the pump wavelength of 1550 nm. The back facet has a broad-band high reflectivity coating of ~90% for both pump and Raman wavelengths. The high reflectivities for the Stokes wavelength reduces the cavity loss so that CW lasing is achievable.

Ring Cavity Laser

Next we review our fully monolithic integrated ring cavity silicon Raman laser [78], which allows dimension scalability and on-chip integration with other silicon photonic components. The ring cavity is constructed from a low-loss rib waveguide forming a racetrack-shaped laser cavity with an integrated p-i-n structure on a single chip (Fig. 27.16). A bus waveguide is connected with the ring cavity via a directional coupler that couples both pump and signal laser light into and out of

the cavity. The coupling ratio depends on input wavelength and polarization and can be varied by changing the gap and/or length of the coupler [79].

27.3.4.3. Hybrid Silicon Laser

Currently, lasers may be integrated with silicon by first separately fabricating individual semiconductor lasers from III–V materials and then attaching them to the silicon photonic die by aligning them one at a time. This has the large drawback that the assembly time/cost increases as the number of lasers increases. A better approach would be to use the same wafer-scale fabrication techniques to manufacture both the silicon components and the laser.

Until an electrically pumped silicon laser is demonstrated, III-V semiconductor materials will continue to be necessary to generate light and produce communication lasers. One must then determine how these optical sources can be integrated cost-effectively onto the silicon photonic chip. Figure 27.17 shows a simple schematic of different approaches of bringing a III-V light source onto a silicon photonic chip.

Perhaps the most straightforward way to introduce light into a silicon photonic device is with a fiber-coupled laser. The fiber delivering the laser's energy could be connected and passively aligned to the silicon die. Once into the silicon, the laser light could be split into separate channels, each of which could be modulated before being recombined and transmitted. This is directly analogous to the way an electrical supply powers an integrated circuit. The advantage of this approach is the flexibility in choice of the laser source, because many different fiber-coupled sources are commercially available already. Another advantage is that the laser is separate from the silicon device, allowing better thermal management of the photonic system. The disadvantage of this approach is its relatively high cost for the laser source, the size, and the packaging complexity.

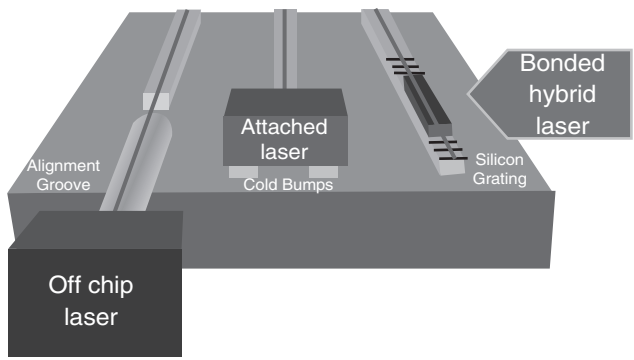


Figure 27.17 Three approaches to integrating a laser into a silicon-device are (left to right) a hybrid laser, an on-chip laser, and an off-chip laser.

A better option for coupling light into a silicon photonic chip is to bond, or solder, individual low-cost lasers directly to the silicon chip, using mechanical stops to align the lasers. The issue here is that submicron alignment accuracy is needed when coupling to small, single-mode waveguides, and this is hard to achieve in a high-volume manufacturing environment with today's pick-and-place tools. Moreover, because each laser needs to be placed and aligned individually, the assembly time increases linearly with the number of lasers required. As the number of lasers per chip increases, this approach becomes prohibitively inefficient and costly.

The third and even better approach is the hybrid silicon laser developed recently from collaboration between the University of California at Santa Barbara and Intel Corporation [8]. This approach allows one to overcome the many problems posed by the previous integration strategies, by using a volume manufacturable wafer-bonding process. In this approach, an unprocessed III-V wafer is directly bonded to a silicon wafer patterned with optical waveguides. Hybrid lasers are then formed using standard planar-fabrication techniques. Because the silicon waveguides are patterned before laser fabrication, no alignment is needed between the unpatterned III-V wafer and the patterned silicon waveguide wafer. With this approach, we believe that tens—if not hundreds—of lasers can be fabricated simultaneously from a single bonding step. Fabrication may be done at the wafer-, partial-wafer, or die-level, depending on the exact needs and economics of the device being fabricated.

The huge potential in silicon photonics techology (miniaturization, integration with microelectronics, power reduction, optical interconnects, and significant cost reduction) has started to be explored and it is very likely that during the next decade silicon will become the main photonic material for the benefit of diverse areas of applications. We are at the beginning of this new techology adoption and while there is sigificant work ahead the future of silicon photonics to revolutionize communication, computing, and other fields looks very promising.

ACKNOWLEDGMENT

The author would like to thank all the members of the Silicon Photonics groups at Intel (Fab 8 group at Jerusalem and Photonics technology labs—PTL at Santa Clara) for all their valuable contributions to this work and for stimulating discussions.

REFERENCES

1. Liao, L., D. Samara-Rubio, M. Morse, A. Liu, H. Hodge, D. Rubin, U. D. Keil, and T. Franck. 2005. High-speed silicon Mach-Zehnder modulator. *Opt. Express* 13:3129–3135.
2. Gardes, F. Y., G. T. Reed, N. G. Emerson, and C. E. Png. 2006. A sub-micron depletion-type photonic modulator in silicon on insulator. *Optics Express* 13:8845–8853.

3. Liu, J., J. Michel, W. Giziewicz, D. Pan, K. Wada, D. Cannon, S. Jongthammanurak, D. Danielson, L. C. Kimerling, J. Chen, F. O. Ilday, F. X. Kartner, and J. Yasaitis. 2005. High performance, tensile-strained Ge p-i-n photodetectors on a Si platform. *Appl. Phys. Lett* 87:103501–103503.
4. Bogaerts, W., R. Baets, P. Dumon, V. Wiaux, S. Beckx, D. Taillaert, B. Luyssaert, J. Van Campenhout, P. Bienstman, and D. Van Thourhout. 2005. Nanophotonic waveguides in siliconon-insulator fabricated with CMOS technology. *J. Lightw. Technol.* 23(1):401–412.
5. Almeida, V. R., R. R. Panepucci, and M. Lipson. 2003, August. Nanotaper for compact mode conversion. *Opt. Lett.* 28(15):1302–1304.
6. Masanovic, G. Z., G. T. Reed, W. Headley, and B. Timotijevic. 2005, September. A high efficiency input/output coupler for small silicon photonic devices. *Opt. Express*, 13(19): 7374–7379.
7. Pavesi, L., S. Gaponenko, and L. Dal Negro. 2003. *Towards the first silicon laser*. NATO science series, Kluwer Academic Publications.
8. Fang A. W., et al. 2006. Electricaly pumped hybrid AlGaInAs-silicon evanescent laser. *Opt. Express*, 14:9203–9210.
9. Boyraz O., and B. Jalali. 2004. Demonstration of a silicon Raman laser. *Opt. Express* 12:5269–5273.
10. Rong, H., et al. 2005. A continuous-wave Raman silicon laser. *Nature* 433:725–728.
11. Rong, H., Y.-H. Kuo, S. Xu, A. Liu, R. Jones, M. Paniccia, O. Cohen, and O. Raday. 2006. Monolithic integrated Raman silicon laser. *Opt. Express* 14:6705–6712.
12. Barwicz, T., et al. 2007. Silicon photonics for compact, energy-efficient interconnects [Invited]. *J. of Opt. Net.* 6(1):63–73.
13. Paniccia, M., M. Morse, and M. Salib. 2004. Integrated photonics. In *Silicon photonics*, eds. L. Pavesi and D. J. Lockwood. Berlin Heidelberg: Springer, pp. 51–88.
14. Kitoh, T., N. Takato, M. Yasu, and M. Kawachi. 1995. Bending loss reduction in silica-based waveguides by using lateral offsets. *J. Lightwave Technol.* 13:555–562.
15. Vlasov, Y. A., and S. J. McNab. 2004. Losses in single-mode silicon-on-insulator strip waveguides and bends. *Opt. Express*, 12:1622–1631.
16. Cao, G. B., L. J. Dai, Y. J. Wang, J. Jiang, H. Yang, and F. Zhang. 2005. Compact integrated star coupler on silicon-on-insulator. *IEEE Photon. Technol. Lett.* 17:2616–2618.
17. Day, I., I. Evans, A. Knighs, F. Hopper, S. Roberts, J. Johnston, S. Day, I. Luff, H. Tsang, and M. Asghari. 2003. Tapered silicon waveguides for low insertion loss highly-efficient high-speed electronic variable optical attenuators. *OFC* 1:249.
18. Reed, G. T., and A. P. Knights. 2004. *Silicon photonics, an introduction*. Haboken, N.J.: John Wiley.
19. Doylend, J. K., and A. P. Knights. 2006. Design and simulation of an integrated fiber-to-chip coupler for SOI waveguides. IEEE *J. of Select. Topic. in Quant. Elect.* 12:1363–1370.
20. Gunn, C. 2006. CMOS photonics for high-speed interconnects. *IEEE Micro* 26:58–66.
21. Izhaky, N., and A. Hardy. 1999. Analysis of grating-assisted backward coupling employing the unified coupled-mode formalism. *J. Opt. Soc. Am. A* 16:1303–1311.
22. Soldano, L. B., and E. C. M. Pennings. 1995. Optical multimode interference devices based on self-imaging: principles and applications. *J. Lightwave Tech.* 13:615–627.
23. Bryngdahl, O. 1973. Image formation using self-imaging techniques. *J. Opt. Soc. Amer.* 63:416–419.
24. Ulrich, R. 1975. Image formation by phase coincidences in optical waveguides. *Opt. Commun.* 12:259–264.
25. Fischer, U., T. Zinke, and K. Petermann. Integrated optical waveguide switches in SOI. 1995. *Proc. 1995 IEEE International SOI Conference*: 141–142.
26. Trinh, P. D., S. Yegnanarayanan, F. Coppinger, and B. Jalali. Compact multimode interference couplers in silicon-on-insulator technology. 1997. *Proc. CLEO'97*:441.

27. Dai, D., S. Liu, S. He, and Q. Zhou. 2002. Optimal design of an MMI coupler for broadening the spectral response of an AWG demultiplexer. *J. Lightwave Tech.* 20:1957–1961.
28. Jiang, X., X. Li, H. Zhou, J. Yang, M. Wang, Y. Wu, and S. Ishikawa. 2005. Compact variable optical attenuator based on multimode interference coupler. *IEEE Photon. Tech. Lett.* 17:2361–2363.
29. Burns, W. K., et al. 1976. *Appl. Phys. Lett.* 29:790.
30. Sasaki, H., and R. M. De La Rue. 1976. *Elect. Lett.* 12:459.
31. Silberberg, Y., et al. 1987. Digital optical switch. *Appl. Phys. Lett.* 51:1230–1232.
32. Rickman, A. G., and G. T. Reed. 1994. Silicon-on-insulator optical rib waveguides: loss, mode characteristics, bends, and y-junctions. *IEE Proc-Optoelect.* 141:391–393.
33. Tang, C. K., A. K. Kewell, G. T. Reed, A. G. Rickman, and F. Namavar. 1996. Development of library of low-loss silicon-on-insulator optoelectronic devices. *IEE Proc-Optoelect.* 143:312–315.
34. Dumon, P., et al. 2004. Basic photonic wire components in silicon-on-insulator. *IEEE International Conference on Group IV Photonics*: 1–3.
35. Liu, Y. L., et al. 2005. Silicon 1 · 2 digital optical switch using plasma dispersion. *Elect. Lett.* 30:130–131.
36. Borel, P. I., et al. 2005. Topology optimized broadband photonic crystal Y-splitter. *Elect. Lett.* 41: 69–71.
37. Izhaky, N., et al. 2001. Intelligent switches of integrated lightwave circuits with core telecommunication functions. *Proc. SPIE,* 4284:54–63.
38. Tang, C. K., G. T. Reed, A. J. Walton, and A. G. Rickman. 1994. Low loss single mode optical phase modulator in SIMOX material. *J. Lightwave Tech.* 12:1394–1400.
39. Atta, R. M. H., et al. 2003. Fabrication of a highly efficient optical modulator based on siliconon-insulator. *Proc. SPIE* 5116:495–500.
40. Liu, A., R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia. 2004. A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor. *Nature* 427:615–618.
41. Liu, A., L. Liao, D. Rubin, H. Nguyen, B. Ciftcioglu, Y. Chetrit, N. Izhaky, and M. Paniccia. 2007. High-speed optical modulation based on carrier depletion in a silicon waveguide. *Opt. Express* 15:660–668.
42. Espinola, R. L., et al. 2003. Fast and low-power thermooptic switch on thin silicon-on-insulator. *IEEE Photon. Tech. Lett.* 15:1366–1368.
43. Jalali, B., et al. 1996. Guided-wave optics in silicon-on-insulator technology. *IEE Proc.-Optoelectron.* 143:307–311.
44. Liu, Y., et al. 2004. Silicon-on-insulator waveguide Mach-Zehnder wavelength combiner. *IEEE International Conference on Group IV Photonics*: 536–537.
45. Wang, Z., et al. 2005. Rearrangeable nonblocking thermo-optic 4 · 4 switching matrix in siliconon-insulator. *IEE Proc.-Optoelectron.* 152:160–162.
46. Li, Y., J. Yu, and S. Chen. 2005. Rearrangeable nonblocking SOI waveguide thermooptic 4 · 4 switch Matrix with low insertion loss and fast response. *IEEE Photon. Tech. Lett.* 17:1641–1643.
47. Othonos, A., and K. Kalli. 1999. *Fiber Bragg gratings: fundamentals and applications in telecommunicationsand sensing,* London: Artech House.
48. Subramanian, V. R., R. G. DeCorby, J. N. McMullin, C. J. Haugen, and M. Belov. 2002. Fabrication of aperiodic gratings on Silicon-On-Insulator (SOI) rib waveguides using e-beam Lithography *Proc. SPIE* 4654:45–53.
49. Murphy, T. E., J. T. Hastings, and H. I. Smith. 2001. Fabrication and characterization of narrowband bragg-refl ection filters in silicon-on-insulator ridge waveguides. *J. Lightwave Technol.* 19:1938–1942.

50. Chu, T., H. Yamada, S. Ishida, and Y. Arakawa. 2005. Reconfigurable optical add-drop multiplexer based on silicon nano-wire waveguides. *ECOC Proceedings* 2:245–246.
51. Moss, D. J., V. G. Ta'eed, B. J. Eggleton, D. Freeman, S. Madden, M. Samoc, B. Luther-Davies, S. Janz, and D.-X. Xu. 2004. Bragg gratings in silicon-on-insulator waveguides by focused ion beam milling. *Appl. Phys. Lett.* 85:4860–4862.
52. Lee, K. K., D. R. Lim, and L. C. Kimerling. 2001. Fabrication of ultralow-loss Si/SiO2 waveguides by roughness reduction. *Opt. Lett.* 26:1888–1890.
53. Sparacin, D. K., S. J. Spector, and L. C. Kimerling. 2005. Silicon waveguide sidewall smoothing by wet chemical oxidation. *J. Lightwave Technol.* 23:2455–2461.
54. Vorckel, A., M. Monster, P. Haring Bolivar, H. Kurz, and W. Henschel. 2003. Fabrication and characterization of ultra compact silicon-on-insulator micro-ring resonator add-drop-multiplexer for dense wavelength division multiplexing. *Conference on Lasers and Electro-Optics*, pp. 866–867.
55. Dumon, P., W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, J. Van Campenhout, D. Taillaert, B. Luyssaert, P. Bienstman, D. Van Thourhout, and R. Baets. 2004. Low-loss SOI photonic wires and ring resonators fabricated with deep UV lithography. *IEEE Photon. Technol. Lett.* 16:1328–1330.
56. Tsuchizawa, T., K. Yamada, H. Fukuda, T. Watanabe, J. Takahashi, M. Takahashi, T. Shoji, E. Tamechika, S. Itabashi, and H. Morita. 2005. Microphotonics devices based on silicon microfabrication technology. *IEEE J. Selected Topics in Quantum Electron.* 11:232–240.
57. Kiyat, I., A. Aydinli, and N. Dagli. 2005. High-Q silicon-on-insulator optical rib waveguide racetrack resonators. *Opt. Express* 13:1900–1905.
58. Headley, W. R., G. T. Reed, S. Howe, A. Liu, and M. Paniccia. 2004. Polarization-independent optical racetrack resonators using rib waveguides on silicon-on-insulator. *Appl. Phys. Lett.* 85:5523–5525.
59. Xu, Q., B. Schmidt, S. Pradhan, and M. Lipson. 2005. Micrometre-scale silicon electro-optic modulator. *Nature* 435:325–327.
60. Xia, F., L. Sekaric, and Y. Vlasov. 2007. Ultra compact optical buffers on a silicon chip. *Nature Photonics* 1(1):65–72.
61. Soref, R. A., and P. J. Lorenzo. 1986. All-silicon active and passive guided-wave components for $\lambda = 1.3$ and 1.6 μm. *IEEE J. Quantum Electron.* QE-22:873–879.
62. Soref, R. A., and B. R. Bennett. 1987. Electrooptical effects in silicon. *IEEE J. Quantum Electron.* QE-23:123–129.
63. Noguchi, K., O. Mitomi, and H. Miyazawa. 1998. Millimeter-wave Ti:LiNbO3 optical modulators. *J. Lightwave Technol.* 16:615–619.
64. Tsuzuki, K., T. Ishibashi, T. Ito, S. Oku, Y. Shibata, T. Ito, R. Iga, Y. Kondo, and Y. Tohmori. 2005. A 40-Gb/s InGaAlAs-InAlAs MQW n-i-n Mach-Zehnder modulator with a drive voltage of 2.3 V. *IEEE Photon. Technol. Lett.* 17:46–48.
65. Kimukin, I., N. Biyikli, B. Butun, O. Aytur, M. S. Ünlü, and E. Ozbay. 2002. InGaAs-based high-performance p-i-n photodiodes. *IEEE Photon. Technol. Lett.* 14:366–368.
66. Iacona F., et al. 2000. Correlation between luminescence and structural properties of Si Nanocrystals. *J. of Appl. Phys.* 87:1295–1303.
67. Jutzi M., M. Berroth, G. Wohl, M. Oehme, and E. Kasper, 2005. Ge-on-Si Vertical Incidence Photodiodes With 39 GHz Bandwidth. *IEEE Photon. Technol. Lett.* 17:1510–1512.
68. Cullis, A. G., et al. 1997. The structural and luminescence properties of porous silicon. *J. of Appl. Phys.* 82:909–965.
69. Polman, A., G. N. van den Hoven, J. S. Custer, J. H. Shin, R. Serna, and P. F. A. Alkemade. 1995. February. Erbium in crystal silicon: Optical activation, excitation, and concentration limits. *J. Appl. Phys.* 77, no. 3:1256–1262.

70. Franzò, G., S. Coffa, F. Priolo, and C. Spinella. 1997, March. Mechanism and performance of forward and reverse bias electroluminescence at 1.54 μm from Er-doped Si diodes. *J. Appl. Phys.* 81, no. 6:2784–2793.
71. Castagna, M. E., et al. 2003. Si-based materials and devices for light emission in silicon. *Physica E* 16:547–553.
72. Liang, T. K., and H. K. Tsang. 2004. Role of free carriers from two-photon absorption in Raman amplification in silicon-on-insulator waveguides. *Appl. Phys. Lett.* 84:2745–2747.
73. Rong, H., et al. 2004. Raman gain and nonlinear optical absorption measurement in a low loss silicon waveguide. *Appl. Phys. Lett.* 85:2196–2198.
74. Liu, A., H. Rong, M. Paniccia, O. Cohen, and D. Hak. 2004. Net optical gain in a low loss silicon-on-insulator waveguide by stimulated Raman scattering. *Opt. Express* 12:4261–4267.
75. Raghunathan, V., O. Boyraz, and B. Jalali. 2005, May. "20 dB On-off Raman amplification in silicon waveguides." CLEO 2005, CMU1, Baltimore, Md.
76. Jones, R., et al. 2005. Net continuous wave optical gain in a low loss silicon-on-insulator waveguide by stimulated Raman scattering. *Opt. Express* 13:519–525.
77. Sih, V., S. Xu, Y.-h. Kuo, H. Rong, M. Paniccia, O. Cohen, and O. Raday. 2007. Raman amplification of 40 Gb/s data in low-loss silicon waveguides. *Opt. Express* 15:357–362.
78. Rong, H., Y.-H. Kuo, S. Xu, A. Liu, R. Jones, M. Paniccia, O. Cohen, and O. Raday. 2006. Monolithic integrated Raman silicon laser. *Opt. Express* 14:6705–6712.
79. Headley, W. R., G. T. Reed, S. Howe, A. Liu, and M. Paniccia. 2004. Polarization-independent optical racetrack resonators using rib waveguides on silicon-on-insulator. *Appl. Phys. Lett.* 85:5523–5525.

This page intentionally left blank

28

Nanophotonics and Nanofibers

Limin Tong
State Key Laboratory of Modern Optical Instrumentation, and Department of Optical Engineering, Zhejiang University, Hangzhou 310027, China

Eric Mazur
Department of Physics and School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA

28.1. INTRODUCTION

Nanophotonics is a fusion of photonics and nanotechnology, and is defined as nanoscale optical science and technology that includes nanoscale confinement of radiation, nanoscale confinement of matter, and nanoscale photoprocesses for nanofabrication [1–3]. While photonics has been widely used for fiber-optic data communication for decades, the application of nanotechnology for optical communication is an emerging technology. The basic motivation for incorporating photonics with nanotechnology is spurred by the requirement of increased integration of photonic devices for a variety of applications such as higher data transmission rates, faster response, lower energy consumption, and denser data storage [2]. For example, to reach an optical data transmission rate as high as 10 Tb/s, the size of photonic matrix switching devices should be reduced to 100-nm scale [4].

Historically, the nanotechnology was first proposed by Richard Feynman in his famous talk "There's Plenty of Room at the Bottom"—in the 1959 Annual Meeting of the American Physical Society [5] and has been thriving since the 1980s. Now, incorporated with physics, chemistry, materials, electronics, and biology, nanotechnology has spawned a number of new multidisciplinary areas. Compared with many other nano-incorporated fields such as nanoelectronics, nanophotonics is relatively new. In the past century, the application of nanotechnology in optics or photonics has usually been associated with the near-field scanning optical microscope (NSOM), with emphasis on near-field optics [6–8]. Recently, the rapid development of nanotechnology in photonics, together with

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

the emerging new frontiers such as surface plasmonics, optical antennas, negative-index metamaterials, and nanofibers in subwavelength optics [9], has greatly broadened the topic to cover a wider scope of light-matter interactions on the nanoscale, and brought new opportunities for future photonic applications, including optical communications, sensing, computing, and storage.

To illustrate this nano-induced potential for future fiber-optic data communication, this chapter provides brief introductions to nanophotonics and nanofibers, with the hope of providing insight to stimulate deeper understanding of this new frontier.

28.2. NANOPHOTONICS

The basic foundation of nanophotonics involves light–matter interaction on the nanometer scale, which is usually extended to subwavelength scale within the optical domain [3]. Physically speaking, in almost all nanophotonic processes concerned, light interacts with electrons (or, equivalently, electrons such as holes and excitons) inside the concerned materials. Therefore, theoretically most of the nanophotonic processes can be fully described by the Maxwell equations and the Schrödinger equation. However, the mesoscopic feature size of the nanophotonic structure, ranging from the atomic scale to the wavelength of the light, makes it difficult or impossible to precisely treat these systems the same way as individual constituents (e.g., a single or cluster of atoms) or large ensembles (e.g., a bulk material) without the perceptible or sometimes predominant optical near-field, surface, and quantum effects.

As has been mentioned, nanophotonics covers a broader area ranging from nanoscale optical engineering and quantum optics to biotechnology. Within the scope of this book, or more specifically on the technological side of optical communication, we are attempting to provide a conceptual introduction to some relative topics, including evanescent waves, surface plasmonics, and quantum confined effects. While we cannot claim comprehensive coverage of all relative topics, we hope that this work will serve as a valuable reference for the readers.

28.2.1. Optical Near-Field and Evanescent Waves

The near-field can be defined as the extension outside a given material of the field existing inside this material [10]. Generally, the amplitude of near-field maintains an evident value in the vicinity of the material boundary but decays very rapidly along the direction perpendicular to the interface, giving rise to the so-called evanescent wave character of the near-field. In optics, the field is manifested by the electromagnetic waves, and near-field optics deals with phenomena involving evanescent electromagnetic waves. In most cases, when the size of the structure goes to a subwavelength or a smaller scale, the effect

of evanescent field becomes significant, making it one of the primary topics of nanophotonics.

Historically, the research of optical evanescent waves can be traced back four hundred years [11], when Sir Isaac Newton investigated the frustrated total reflection of a prism. As schematically illustrated in Fig. 28.1, in one of his experiments, Newton placed a prism against a lens and observed that the light intensity transferred onto the lens was located on an area of the lens much larger than the point of contact. The result indicates that, when the air gap between the lens and the prism is very small, light confined by the prism tunnels through the gap and leaks into the lens, which is attributed to the existence of evanescent waves in the vicinity of the reflection surface.

Mathematically, evanescent fields in total reflection can be described using Maxwell's equations and boundary conditions of the electromagnetic fields. As a typical example, shown in Fig. 28.2 is reflection of light on a plane interface of two dielectric media with refractive indices of n_1 and n_2, respectively. Assume the incident light with the wavelength of λ to be a plane wave with an electric component $E_1 = E_{10}e^{i(k_{1x}x+k_{1z}z-\omega t)}$. When the incident angle θ_1 is larger than the critical angle θ_c [$\theta_c = \sin^{-1}(n_2/n_1)$], total reflection occurs with a penetration field ($z > 0$ region) $E_3 = E_{30}e^{i(k_{3x}x+k_{3z}z-\omega t)}$, where

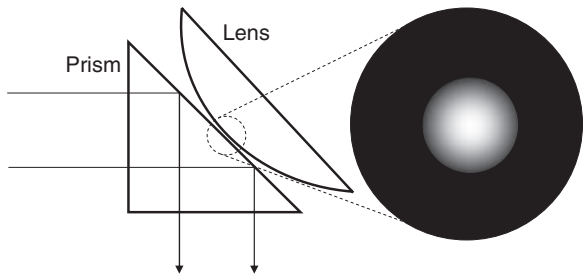


Figure 28.1 Schematic illustration of the frustrated total reflection of a prism.

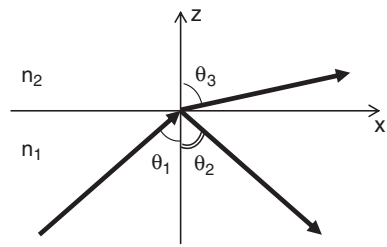


Figure 28.2 Reflection of light on a plane interface.

$$k_{3z} = k_3 \cdot \sqrt{1 - \sin^2(\theta_3)} = \frac{2\pi}{\lambda} \cdot \sqrt{1 - \frac{n_1^2 \sin^2(\theta_1)}{n_2^2}} \tag{28.1}$$

When $\theta_1 > \theta_c$, $\sqrt{1 - \frac{n_1^2 \sin^2(\theta_1)}{n_2^2}}$ is imaginary, Eq. (28.1) can be expressed as

$$k_{3z} = i\kappa = i\frac{2\pi}{\lambda} \cdot \sqrt{n_1^2 \sin^2(\theta_1) - n_2^2} \tag{28.2}$$

where κ is real. The penetration field is then written as

$$E_3 = E_{30} e^{-\kappa z} e^{i(k_{3x}x - \omega t)} \tag{28.3}$$

Therefore, E_3 is an evanescent field and decays exponentially along the z-direction with a penetration depth

$$d = \frac{1}{\kappa} = \frac{\lambda}{2\pi\sqrt{n_1^2 \sin^2(\theta_1) - n_2^2}} \tag{28.4}$$

For reference, with a green light (λ~500 nm) and an angle of incidence of 60°, assume n_1 and n_2 to be 1.5 (e.g., glass) and 1.0 (air); the penetration length d is about 100 nm.

Evanescent-field-induced effects are observed in many macroscopic structures. For example, Goos-Hänchen shifts in standard optical fibers [12]. However, when the feature size of an optical structure goes to subwavelength or nanometer scale, the effect of evanescent field will be dominant. For example, when a He-Ne laser (633-nm wavelength) is guided along a 200-nm-diameter air-clad silica nanofiber, over 90% of the light energy is guided outside the silica core as evanescent waves [13]. The high fraction of evanescent waves may lead to significant enhancement of the optical interaction between closely located structures [14] and may find a variety of nanophotonic applications, including evanescent-field-based optical coupling, sensing, manipulation, excitation, and nonlinear effects.

The characterization of evanescent fields has been greatly facilitated by the invention of the near-field scanning optical microscope (NSOM) [15]. When scanning and picking up signals by a near-field probe that is kept less than 10 nm away from the interface, NSOM can efficiently collect light energy carried by the evanescent field and offer a spatial resolution better than 50 nm. By operating it reversely, NSOM can also be used for launching light into a very small area for highly localized excitation. The capability for picking up evanescent waves and imaging beyond the optical diffraction limit makes the NSOM a powerful tool for both investigation and modification of nanophotonic structures.

28.2.2. Surface Plasmon Resonance

Surface plasmon resonance (SPR) is the collective oscillation of electrons on the interface between a metal and a dielectric medium. This kind of oscillation, highly localized on the surface of the metal-dielectric boundary, decays quickly along the direction perpendicular to the interface. One of the well-known early works on SPR was reported by R. H. Ritchie in 1957 [16], in which the plasma losses by fast electrons in thin metallic films were investigated theoretically. In 1970s, J. Shoenwald et al. experimentally observed the propagation of surface polaritons on a metal-air surface at optical frequencies [17]. Due to the low-loss feature of metals at infrared wavelength, they found that these localized surface waves could propagate over macroscopic distances. Recently, incorporated with nanotechnology, SPR has brought considerable changes in the optical properties of nanostructures, such as enhanced transmission in subwavelength aperture and optical waveguiding in metal nanoparticle chains [18–20].

A schematic illustration of SPR is shown in Fig. 28.3. Assume an interface located at $z = 0$ to be the interface of the two media with permittivities of ε_1 (Medium 1) and ε_2 (Medium 2), respectively. The surface wave propagated along y-direction. To keep the electromagnetic energy around the interface, the boundary conditions and the symmetry of the electric fields (i.e., the z-components of the electric field in the two half spaces should always take opposite signs) require that

$$\varepsilon_1 = -C\varepsilon_2 \tag{28.5}$$

where C is a positive constant.

Equation (28.5) means that for sustaining a surface wave, the permittivities of the two media should take opposite signs. At optical frequency, dielectric (e.g., air) and metal (e.g., silver) are the best choices.

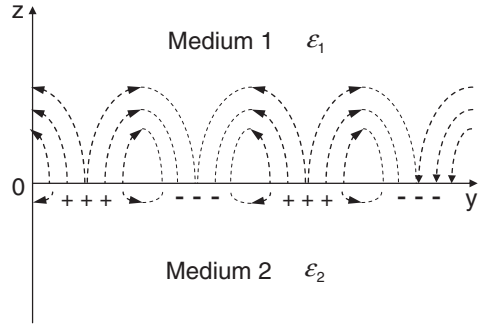


Figure 28.3 Schematic illustration of surface plasmon resonance (SPR).

By solving Maxwell's equations under appropriate boundary conditions, the dispersion relation for these surface waves can be obtained as [20]

$$k_{sp} = k_y = \frac{\omega}{c} \cdot \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}} \tag{28.6}$$

Generally, the momentum of the surface wave (k_{sp}) obtained in Eq. (28.6) is larger than that of the light wave propagating in free space (i.e., in medium 1), indicating the difficulty for direct coupling between the free-space light and the surface waves.

To achieve efficient coupling between the light and the surface wave, there are several approaches for compensation of the momentum mismatch [17], among which the introduction of periodic structures is one of the most effective and practical approaches. As shown in Fig. 28.4, when a light beam is incident on a one-dimensional grating with a period of *H*, the momentum of light in the y-direction is

$$k_y - \frac{2\pi}{\lambda} \sin\theta \pm \frac{2\pi n}{H} \tag{28.7}$$

where θ is the angle of incidence, λ is the wavelength of the light, and *n* is an integer. The second term in the right side of Eq. (28.7) is the additional momentum from the grating, which can be used for compensating the momentum mismatch. To get a high coupling efficiency within the visible or near-infrared spectral range, the grating period *H* should be on a subwavelength or nanometer scale.

The potential of SPR for future optical data processing originates from its capability of utilizing low-dimensional surface waves to provide deep subwavelength or nanoscale confinement, guidance, and switch of light energy [21]. For example, in 2003, S. A. Maier et al. demonstrated light energy transport in metal nanoparticle plasmon waveguides at 570-nm wavelength [19]. The cross section of the particle chain is below 100 nm, and a 0.5-μm propagation length was

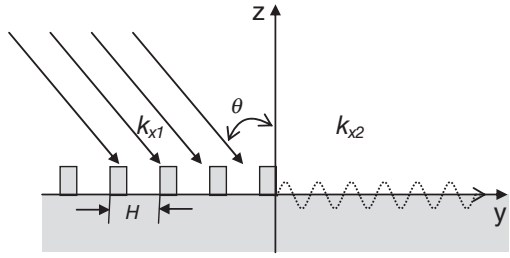


Figure 28.4 Schematic diagram of grating coupling between free-space light and surface plasmon waves.

observed. More recently, SPR-based photonic devices such as optical interferometers, filters, and resonators have been demonstrated [22, 23].

28.2.3. Quantum Confined Effects

Generally, the wave-like nature of the electrons becomes obvious when the dimension of the confinement approaches or becomes smaller than the de Broglie wavelength of the electron. In most cases, this kind of quantum confinement is obtained or enhanced by size reduction of the nanostructure, and can be used to control and sometimes introduce new optical properties of the nanostructures.

Nanostructurized inorganic semiconductor is one of the most important and widely used materials for manifestation of the quantum confinement effect, which provides an added dimension to the highly active area of "bandgap engineering" of the semiconductor bandgap [3]. So far, the quantum-confined structures of many different types of semiconductors have been demonstrated, such as quantum wells, quantum wires, quantum dots, and superlattices.

For a brief introduction, Fig. 28.5 shows a schematic illustration of a quantum well, in which a layer of a small bandgap semiconductor with thickness on the nanometer scale is sandwiched between two layers of a wider bandgap semiconductor. The layered structure provides the potential for confining the electrons and the holes, resulting in quantized energy levels of these charge carriers when their kinetic energies are lower than the potential provided by the well. In a simplest case that the potential well is a one-dimension box with infinite height, the energy of the electron in the conduction band is given as [3]

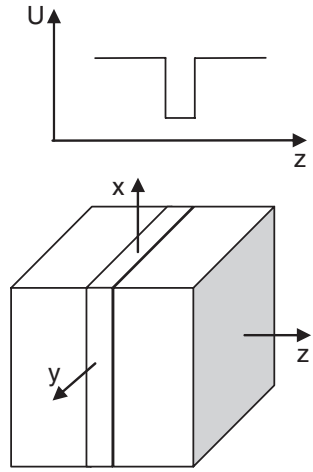


Figure 28.5 Schematic diagram of a quantum well.

$$E_{n,k_x,k_y} = E_C + \frac{n^2h^2}{8m_e^*l^2} + \frac{h^2(k_x^2 + k_y^2)}{8\pi^2 m_e^*} \tag{28.8}$$

where h is the Planck constant, l is width of the well, E_c is the lowest energy of the electron at the bottom of the conduction band, m_e^* is the effective mass of the electron, n is an integer larger than zero, and k_x and k_y are momentum of the electron in the x- and y-direction. The second term on the right-hand side represents the quantized energy. Since the minimum value of n is 1, the quantization increases the energy of the electron in the conduction band. For the hole in the valence band, the quantization decreases its energy because of its positive charge. Therefore, the bandgap of the semiconductor in the quantum well is increased and becomes size-dependent.

Besides the change of the energy, another important feature of quantum confined structure is the modification of the density of the states (DOS), which is also directly associated with the optical properties of the nanostructure. For example, in bulk materials, the DOS of an electron is zero at the bottom of the conduction band and is continuously increased with increasing energy, while in a quantum dot, the DOS is a series of discontinuous peaks. That means DOS has discrete values only at some allowed energy, which is similar to the energy level of an atom and makes the quantum dot an "artificial atom".

The quantum effects in nanostructures bring new opportunities for nanophotonic devices. For example, compared with heterostructure lasers without tight quantum confinement, quantum well lasers offer lower threshold, narrow spectral gain, and much higher modulation frequency [24]. Another example is the quantum cascade laser that relies on the intraband transition of electrons between subbands generated by quantum confinement in quantum well superlattices [25], which provides an excellent solution for the infrared laser source. More recently, nanophotonic NOT gates using near-field optically coupled quantum dots [26] have been experimentally demonstrated.

28.3. NANOFIBERS

Glass fiber is no doubt the backbone of optic data communication and optical networking [27]. Recent advances in nanophotonics have spurred efforts for the miniaturization of optical fibers and fiber-optic devices. When the diameter of a fiber goes well below 1 micrometer, it is called a nanofiber. The nanofiber is usually operated with large index contrast between the glass core and the cladding atmosphere (e.g., air). The small diameter of a nanofiber and the large core-cladding index contrast yields a number of interesting optical properties such as tight optical confinement, large evanescent fields, strong field enhancement, and large waveguide dispersions. An important motivation for nanoscale fiber optics

is its potential usefulness as building blocks in future micro- or nanophotonic devices for optical communications, sensors, and other purposes.

28.3.1. Nanofiber Fabrication

The optical nanofibers introduced here are fabricated by high-temperature taper drawing of standard optical fibers or bulk glasses, which is a top-down process that permits the fabrication of nanofibers with high uniformities [28–30]. Figure 28.6 shows typical electron microscope images of silica nanofibers drawn from standard optical fibers. For example, Fig. 28.6(a) is a scanning electron microscope (SEM) image of a 50-nm-diameter nanofiber; the high uniformity of its diameter is clearly seen. In Fig. 28.6(b), a 260-nm-diameter silica nanofiber is coiled up to show its length. Figure 28.6(c) is an SEM image of silica nanofibers with diameters ranging from 230 to 660 nm, all with high uniformities. To investigate the sidewall roughness, Fig. 28.6(d) shows a higher-magnification transmission electron microscope (TEM) image of a 330-nm-diameter silica nanofiber. No visible irregularity is found.

Besides the silica fiber, a variety of other types of glasses (e.g., phosphate, fluoride, and tellurite glasses) have also been drawn into

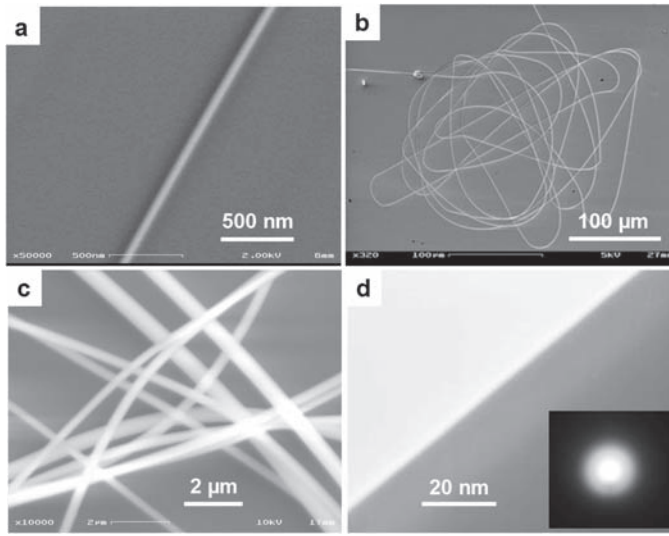


Figure 28.6 Electron microscope images of silica nanofibers taper-drawn from standard optical fibers. (a) SEM image of a 50-nm-diameter nanofiber. (b) SEM image of a coiled 260-nm-diameter nanofiber with a total length of about 4 mm. (c) SEM image of nanofibers with diameters ranging from 230 to 660 nm. (d) TEM image of the sidewall of a 330-nm-diameter nanofiber; the electron diffraction pattern (inset) shows that the fiber is amorphous. (Adapted from Ref. 28.)

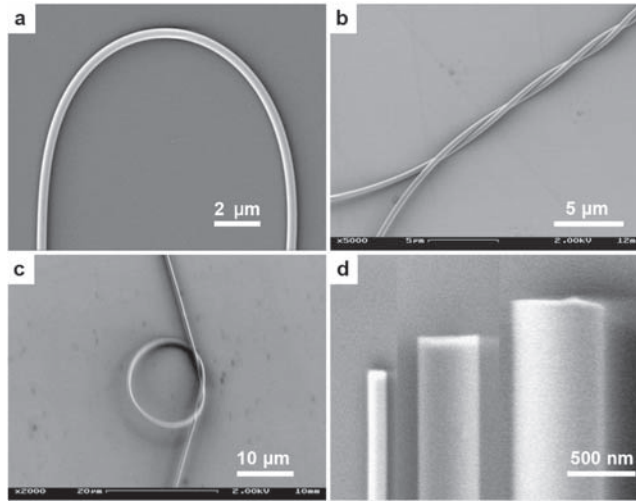


Figure 28.7 SEM images of silica nanofibers patterned with micromanipulation. (a) A bend of 5-μm radius formed with a 410-nm-diameter nanofiber. (b) Two twisted 400-nm-diameter nanofibers. (c) A 14.5-μm-diameter ring assembled using a 520-nm-diameter nanofiber. (d) Flat end faces of a 140-, 420- and 680-nm-diameter nanofibers obtained using a bend-to-fracture process.

highly uniform nanofibers, with diameters down to 50 nm and lengths up to millimeters [30].

Because of their high uniformities, taper-drawn nanofibers also show favorable mechanical properties for micromanipulation and processing. Shown in Fig. 28.7 are SEM images of silica nanofibers patterned with micromanipulation, in which Fig. 28.7(a) is a bend of 5-μm radius formed with a 410-nm diameter nanofiber, Fig. 28.7(b) is two twisted 400-nm-diameter nanofibers, Fig. 28.7(c) shows a 14.5-μm diameter ring assembled with a 520-nm-diameter nanofiber, and Fig. 28.7(d) shows flat end faces of a 140-, 420- and 680-nm diameter nanofibers obtained using a bend-to-fracture process [31]. The high flexibilities of the nanofibers shown in Fig. 28.7 are favorable for their practical applications in micro- and nanophotonic devices.

28.3.2. Optical Waveguiding Properties

Because of the cylindrical symmetry, the guiding behavior of an optical nanofiber can be obtained using analytical solutions of Maxwell's equations [13, 32]. Normalized propagation constants β/k_0 (also known as the effective index; here β is the propagation constant, and $k_0 = 2\pi/\lambda_0$) of the first four modes of air-clad silica nanofiber are shown in Fig. 28.8(a). When the normalized fiber diameter D/λ_0 falls below 0.73 (the dotted line indicates the cut-off condition for single mode), the

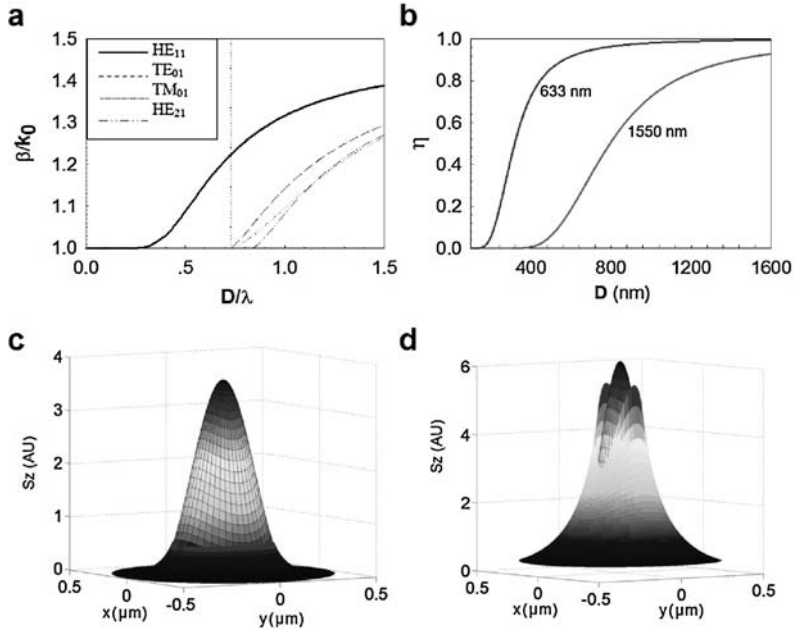


Figure 28.8 Optical wave-guiding properties of air-clad silica nanofibers. (a) Normalized propagation constants β/k_0 of the first four modes. The dotted line indicates the cutoff condition. (b) Fractional power of HE_{11} mode inside the core at 633- and 1550-nm wavelengths. (c) and (d) Z-component Poynting vector in 457- and 229-nm-diameter fibers at 633-nm wavelength, respectively. Mesh, inside silica core; gradient, outside the core.

fiber is a single-mode waveguide. Fig. 28.8(b) shows the fractional power of the fundamental mode (HE_{11}) inside the core at 633- and 1550-nm wavelengths, respectively. For example, at the single-mode cutoff diameter (D_{SM}), more than 80% of the light energy is guided inside the fiber, demonstrating its tight-confinement ability. When the fiber diameter reduces to 0.5 D_{SM}, for example, 229 nm at 633-nm wavelength, about 86% of light power propagates outside the fiber as evanescent waves. For comparison, Figs. 28.8(c) and 28.8(d) give the Poynting vector (along the propagation direction z) of the fundamental mode at 633-nm wavelength with fiber diameters of 457 nm (D_{SM}) and 229 nm (0.5 D_{SM}), respectively. Tight confinement is helpful to reduce the bending loss in a sharp bend; weak confinement, on the other hand, facilitates the light coupling from one wire to another. These favorable guiding properties make the nanofiber promising for building evanescent-coupling-based devices with ultra-compact sizes.

Experimentally, for a nanofiber that is connected to a standard fiber through the tapering region, light guided in the standard fiber can be directly squeezed into the nanofiber through the tapering region. For a nanofiber with both ends

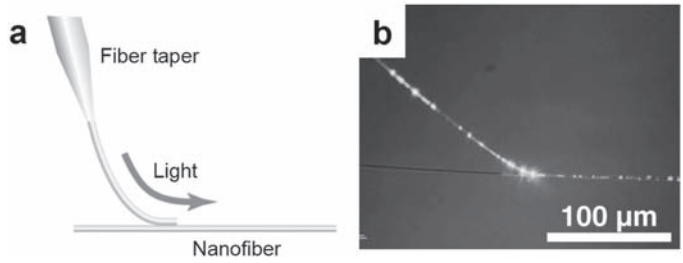


Figure 28.9 Launching light into an optical nanofiber. (a) Schematic diagram for launching light into a nanofiber using the evanescent coupling method. (b) Optical microscope image of a 390-nm-diameter silica nanofiber coupling light into a 450-nm-diameter silica nanofiber. (Adapted from Ref. [28].)

freestanding, light can be launched using an evanescent coupling method. As shown in Fig. 28.9, light is first sent into the core of a standard fiber that is tapered down to a nanofiber; the nanotaper is then used to evanescently couple the light into the specified nanofiber by overlapping the two in parallel. For two nanofibers with the same diameter and refractive index, the coupling efficiency of this kind of evanescent coupling can be higher than 97% [31]; for nanofibers with large difference in refractive index, for example, coupling a 633-nm-wavelength light from a silica nanofiber (1.46 in refractive index) to a tellurite nanofiber (2.0 in index), the coupling efficiency can go up to 90% when the fiber diameter and overlapping length are properly selected [14].

Because of their extraordinary uniformities, taper-drawn nanofibers guide light with low optical losses. Typical loss of taper-drawn glass nanofibers measured at the critical diameter for single-mode operation is lower than 0.1dB/mm [30], with the lowest loss of about 0.001dB/mm measured in silica nanofibers [33], which is much lower than the optical loss of other subwavelength-structures such as surface plasmon waveguides or nanowires fabricated with other methods. The low optical loss, tight optical confinement, strong evanescent field, high uniformity, and mechanical strength make nanofibers promising building blocks for micro- or nanophotonic components.

28.3.3. Device Applications

A variety of micro- or nanoscale photonic components or devices, including optical couplers, interferometers, resonators, and sensors, have been demonstrated using taper-drawn nanofibers [30, 31, 34–39]. Because of their small size, low optical loss, evanescent wave guiding, and mechanical flexibility, these devices show high potential for applications in optical communications and sensors.

Figure 28.10 shows a microscale optical coupler assembled from two tellurite glass nanofibers with diameters of 350 and 450 nm, respectively. When 633-nm-

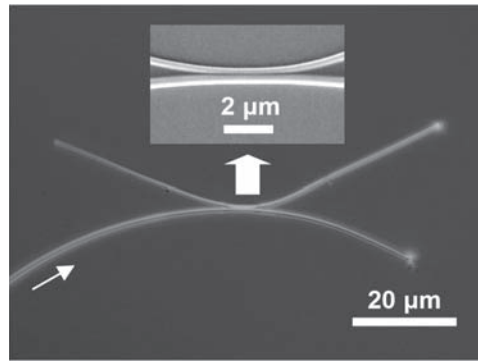


Figure 28.10 Optical micrograph of an optical coupler assembled using two tellurite glass nanofibers (350 and 450 nm in diameter, respectively) on the surface of silica glass. The coupler splits the 633-nm-wavelength light equally. (Adapted from Ref. [30].)

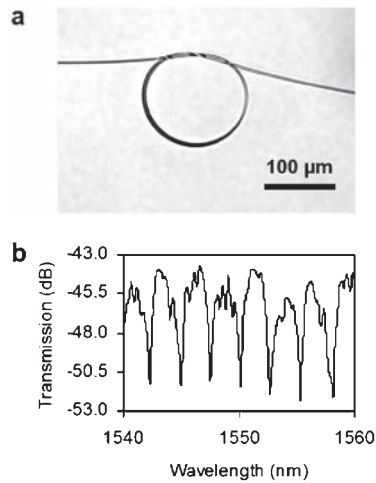


Figure 28.11 Nanofiber-assembled knot resonator. (a) Optical micrograph of a 150-μm-diameter microknot assembled with an 880-nm-diameter silica fiber. (b) Transmission spectrum of the microknot shown in (a).

wavelength light is launched into the bottom left arm, the coupler splits the flow of light in two, working as a 3-dB splitter with almost no excess loss. The overlap length of less than 5 μm is much shorter than the transfer length required by conventional fused couplers made from larger-diameter fiber tapers [40].

When tying a nanofiber into a loop or a knot, a microresonator can be obtained through the evanescent coupling at the joined area. For example, Fig. 28.11(a)

shows a 150-μm-diameter microknot assembled with an 880-nm-diameter fiber. The measured transmittance of this microknot for wavelengths near 1550 nm is shown in Fig. 28.11(b). The spectral response of the knot clearly shows optical resonances with a *Q*-factor higher than 1000. Recently, microcoil/loop resonators with *Q*-factors as high as 95,000 have been realized [41], and a proposed microcoil resonator with self-coupling turns is expected to display a *Q*-factor as high as 10^{10} [42].

Based on microresonators, a series of photonic devices can be realized. Recently, knot-resonator-based add-drop filters and lasers have been experimentally realized using micrometer-diameter silica fibers and rare-earth doped phosphate glass fibers [43, 44]. Rare-earth doped nanofiber ring lasers with much smaller sizes have been theoretically predicted [45], indicating the possibility to develop much more compact devices using nanofibers for optical communications and sensors.

REFERENCES

1. Shen, Y. Z., C. S. Friend, Y. Jiang, D. Jakubczyk, J. Swiatkiewicz, and P. N. Prasad. 2000. Nanophotonics: Interactions, materials, and applications. *J. Phys. Chem. B* 104: 7577–7587.
2. Ohtsu, M., K. Kobayashi, T. Kawazoe, S. Sangu, and T. Yatsui. 2002. Nanophotonics: Design, fabrication, and operation of nanometric devices using optical near fields. *IEEE J. Sel. Top. Quantum Electron.* 8:839–862.
3. Prasad, P. N. 2004. *Nanophotonics*. Hoboken, N.J.: John Wiley & Sons.
4. Kawazoe, T., T. Yatsui, and M. Ohtsu. 2006. Nanophotonics using optical near fields. *J. Non-Cryst. Solids* 352:2492–2495.
5. Feynman, R. P. 1992. There's plenty of room at the bottom. *J. Microelectromech. Syst.* 1:60–66.
6. Paesler, M. A., and P. J. Moyer. 1996. *Near-field optics: Theory, instrumentation, and applications*. New York: John Wiley & Sons.
7. Ohtsu, M., and H. Hori. 1999. *Near-Field nano-optics: From basic principles to nano-fabrication and nano-photonics*. New York: Kluwer Academic.
8. Kawata, S., M. Ohtsu, and M. Irie. 2002. *Nano-Optics*. New York: Springer.
9. Hecht, Jeff. 2005, September. Photonics frontiers: Subwavelength optics come into focus. *Laser Focus World*.
10. Girard, C., C. Joachim, and S. Gauthier. 2000. The physics of the near-field. *Rep. Prog. Phys.* 63:893–938.
11. de Fornel, F. 2001. *Evanescent waves: From Newtonian optics to atomic optics*. Berlin: Springer-Verlag.
12. Saleh, B. E. A., and M. C. Teich. 1991. *Fundamentals of photonics*. New York: John Wiley & Sons.
13. Tong, L. M., J. Y. Lou, and E. Mazur. 2004. Single-mode guiding properties of subwavelength-diameter silica and silicon wire waveguides. *Opt. Express* 12:1025.
14. Huang, K. J., S. Y. Yang, and L. M. Tong. 2007. Modeling of evanescent coupling between two parallel optical nanowires. *Appl. Opt.* 46:1429–1434.
15. Betzig, E., J. K. Trautman, T. D. Harris, J. S. Weiner, and R. S. Kostelak. 1991. Beating the diffraction barrier: Optical microscopy on a nanometer scale. *Science* 251:1468–1470.

16. Ritchie, R. H. 1957. Plasma losses by fast electrons in thin films. *Phys. Rev.* 106:874–881.
17. Shoenwald, J., E. Burstein, and M. Elson. 1973. Propagation of surface polaritons over macroscopic distances at optical frequencies. *Solid State Commun.* 12:185–189.
18. Ebbesen, T. W., H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff. 1998. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature* 391:667–669.
19. Maier, S. A., et al. 2003. Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides. *Nature Mater.* 2:229–232.
20. Barnes, W. L., A. Dereux, and T. W. Ebbesen. 2003. Surface plasmon subwavelength optics. *Nature* 424:824–830.
21. Maier, S. A. 2006. Plasmonics: Metal nanostructures for subwavelength photonic devices. *IEEE J. Sel. Top. Quantum Electron.* 12:1214–1220.
22. Bozhevolnyi, S. I., V. S. Volkov, E. Devaux, J. Y. Laluet, and T. W. Ebbesen. 2006. Channel plasmon subwavelength waveguide components including interferometers and ring resonators. *Nature* 440:508–511.
23. Volkov, V. S., S. I. Bozhevolnyi, E. Devaux, J. Y. Laluet, and T. W. Ebbesen. 2007. Wavelength selective nanophotonic components utilizing channel plasmon polaritons. *Nano Lett.* 7:880–884.
24. Zory, P. S. 1993. *Quantum well lasers*. New York: Academic Press.
25. Capasso, F., C. Gmachl, D. L. Sivco, and A. Y. Cho. 2002, May. Quantum cascade lasers. *Phys. Today* 34–40.
26. Kawazoe, T., K. Kobayashi, K. Akahane, M. Naruse, N. Yamamoto, and M. Ohtsu. 2006. Demonstration of nanophotonic NOT gate using near-field optically coupled quantum dots. *Appl. Phys. B* 84:243–246.
27. Mynbaev, D. K., and L. L. Scheiner. 2001. *Fiber-optic communications technology*, New York: Prentice Hall.
28. Tong, L. M., R. R. Gattass, J. B. Ashcom, S. L. He, J. Y. Lou, M. Y. Shen, I. Maxwell, and E. Mazur. 2003. Subwavelength-diameter silica wires for low-loss optical wave guiding. *Nature* 426:816–819.
29. Brambilla, G., V. Finazzi, and D. J. Richardson. 2004. Ultra-low-loss optical fiber nanotapers. *Opt. Express* 12:2258–2263.
30. Tong, L. M., L. L. Hu, J. J. Zhang, J. R. Qiu, Q. Yang, J. Y. Lou, Y. H. Shen, J. L. He, and Z. Z. Ye. 2006. Photonic nanowires directly drawn from bulk glasses. *Opt. Express* 14:82–87.
31. Tong, L. M., J. Y. Lou, R. R. Gattass, S. L. He, X. W. Chen, L. Liu, and E. Mazur. 2005. Assembly of silica nanowires on silica aerogels for microphotonic devices. *Nano Lett.* 5:259–262.
32. Snyder, A. W., and J. D. Love. 1983. *Optical waveguide theory*. New York: Chapman and Hall.
33. Leon-Saval, S. G., T. A. Birks, W. J. Wadsworth, P. St.J. Russell, and M. W. Mason. 2004. Supercontinuum generation in submicron fibre waveguides. *Opt. Express* 12:2864–2869.
34. Sumetsky, M., Y. Dulashko, and A. Hale. 2004. Fabrication and study of bent and coiled free silica nanowires: Self-coupling microloop optical interferometer. *Opt. Express* 12:3521–3531.
35. Sumetsky, M., Y. Dulashko, J. M. Fini, and A. Hale. 2005. Optical microfiber loop resonator. *Appl. Phys. Lett.* 86:161108.
36. Sumetsky, M. 2005. Uniform coil optical resonator and waveguide: transmission spectrum, eigenmodes, and dispersion relation. *Opt. Express* 13:4331–4340.
37. Lou, J., L. Tong, and Z. Ye. 2005. Modeling of silica nanowires for optical sensing. *Opt. Express* 13:2135–2140.
38. Polynkin, P., A. Polynkin, N. Peyghambarian, and M. Mansuripur. 2005. Evanescent field-based optical fiber sensing device for measuring the refractive index of liquids in microfluidic channels. *Opt. Lett.* 30:1273–1275.

39. Villatoro, J., and D. Monzón-Hernández. 2005. Fast detection of hydrogen with nano fiber tapers coated with ultra thin palladium layers. *Opt. Express* 13:5087–5092.
40. Kakarantzas, G., T. E. Dimmick, T. A. Birks, R. Le Roux, and P. St. J. Russell. 2001. Miniature all-fiber devices based on CO_2 laser microstructuring of tapered fibers. *Opt. Lett.* 26:1137–1139.
41. Sumetsky, M., Y. Dulashko, J. M. Fini, A. Hale, and D. J. DiGiovanni. 2005. Demonstration of the microfiber loop optical resonator. Optical Fiber Communication Conference, Postdeadline papers, Paper PDP10, Anaheim, Calif.
42. Sumetsky, M. 2004. Optical fiber microcoil resonator. *Opt. Express* 12:2303–2316.
43. Jiang, X. S., Y. Chen, G. Vienne, and L. M. Tong. 2007. All-fiber add–drop filters based on microfiber knot resonators. *Opt. Lett.* 32:1710–1712.
44. Jiang, X. S., Q. Yang, G. Vienne, Y. H. Li, L. M. Tong, J. J. Zhang, and L. L. Hu. 2006. Demonstration of microfiber knot laser. *Appl. Phys. Lett.* 89:143513.
45. Li, Y. H., G. Vienne, X. S. Jiang, X. Y. Pan, X. Liu, P. F. Gu, and L. M. Tong. 2006. Modeling rare-earth doped microfiber ring laser. *Opt. Express* 14:7073–7086.

Appendix A

Measurement Conversion Tables

English-to-Metric Conversion Table.

English unit	Multiplied by	Equals metric unit
Inches (in.)	2.54	Centimeters (cm)
Inches (in.)	25.4	Millimeters (mm)
Feet (ft)	0.305	Meters (m)
Miles (mi)	1.61	Kilometers (km)
Fahrenheit (F)	(°F − 32) × 0.556	Celsius (C)
Pounds (1b)	4.45	Newtons (N)

Metric-to-English Conversion Table.

Metric unit	Multiplied by	Equals English unit
Centimeters (cm)	0.39	Inches (in.)
Millimeters (mm)	0.039	Inches (in.)
Meters (m)	3.28	Feet (ft)
Kilometers (km)	0.621	Miles (mi)
Celsius (C)	(°C × 1.8) + 32	Fahrenheit (F)
Newtons (N)	0.225	Pounds (1b)

Absolute Temperature Conversion.

Kelvin (K) = Celsius + 273.15
Celsius = Kelvin − 273.15

Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking
Copyright © 2008, Elsevier Inc. All rights reserved.
ISBN: 978-0-12-374216-2

Area Conversion.

I square meter = 10.76 square feet = 1550 square centimeters
I square kilometer = 0.3861 square miles

Metric Prefixes.

Yotta = 10^{24}	Deci = 10^{-1}
Zetta = 10^{21}	Centi = 10^{-2}
Exa = 10^{18}	Milli = 10^{-3}
Peta = 10^{15}	Micro = 10^{-6}
Tera = 10^{12}	Nano = 10^{-9}
Giga = 10^{9}	Pico = 10^{-12}
Mega = 10^{6}	Femto = 10^{-15}
Kilo = 10^{3}	Atto = 10^{-18}
Hecto = 10^{2}	Zepto = 10^{-21}
Deca = 10^{1}	Yotto = 10^{-24}

Appendix B

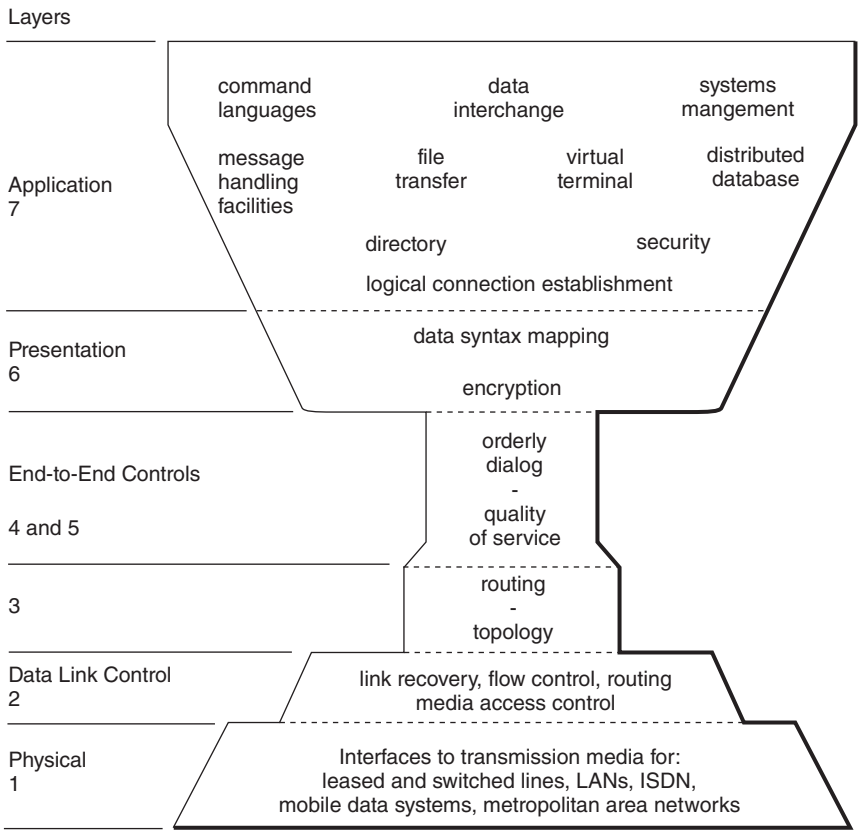
Physical Constants

Speed of light = $c = 2.99792458 \times 10^{8}$ m/s
Boltzmann constant = $k = 1.3801 \times 10^{-23}$ J/K = 8.620×10^{-5} eV/K
Planck's constant = $h = 6.6262 \times 10^{-34}$ J/S
Stephan–Boltzmann constant = $\sigma = 5.6697 \times 10^{-8}$ W/m^2/K^4
Charge of an electron = 1.6×10^{-19} C
Permittivity of free space = 8.849×10^{-12} F/m
Permeability of free space = 1.257×10^{-6} H/m
Impedance of free space = $120\ \pi$ ohms = 377 ohms
Electron volt = 1.602×10^{-19} J

This page intentionally left blank

Appendix C

The 7-Layer OSI Model



The wineglass of layered functions.

This page intentionally left blank

Appendix D

Network Standards and Data Rates

ORGANIZATION OF MAJOR INDUSTRY STANDARDS

The IEEE defines a common path control (802.1) and datalink layer (802.2) for all the following LAN standards. Although FDDI is handled by ANSI, it is also intended to fall under the logical link control of IEEE 802.2 (the relevant IS0 standard is 8802-2).

IEEE LAN Standards

802.3-code sense multiple access/collision detection (CSMNCD) also known as Ethernet. Variants include Fast Ethernet (100BaseX), Gigabit Ethernet (802.32), and 10 Gigabit Ethernet (802.3ae).

802.4-Token bus (TB)

802.5-Token ring (TR)

802.6-Metropolitan area network (MAN) [also sometimes called switched multimegabit data service (SMDS) to which it is related]

802.9-Integrated services digital network (ISDN) to LAN interconnect

802.11-Wireless services up to *5* Mb/s

802.12-100VG AnyLAN standard

802.14-100BaseX (version of Fast Ethernet)

ANSI Standards

Fast Ethernet-ANSI X3.166

Fiber distributed data interface (FDDI): ANSI X3T9.5 (the relevant IS0 standards are IS 9314/12 and DIS 9314/3)

FDDI-ANSI X3.263

Physical layer (PHY)

Physical media dependent (PMD)
Media access control (MAC)
Station management (SMT)
Because the FDDI specification is defined at the physical and data link layers, additional specifications have been approved by ANSI subcommittees to allow for FDDI over single-mode fiber (SMF-PMD), FDDI over copper wire or CDDI, and FDDI over low-cost optics (LC FDDI). A time-division multiplexing approach known as FDDI-II has also been considered.

Serial Byte Command Code Set Architecture (SBCON): ANSI standard X3T11/95-469 (rev.2.2., 1996); follows IBM's enterprise systems connectivity (ESCON) standard as defined in IBM documents SA23-0394 and SA22-7202 (IBM Corporation, Mechanicsburg, PA)

Fibre Channel Standard (FCS)
ANSI X3.230-1994 rev. 4.3, physical and signaling protocol
ANSI X3.272-199x rev. 4.5 (June 1995) Fibre Channel arbitrated loop (FC-AL)

High-Performance Parallel Interface (HIPPI):
Higher speed versions of this interface have been used for technical computing applications. Formerly known as HIPPI 6400 or SuperHIPPI and now known officially as Gigabyte System Network (GSN, a trademark of the High Performance Networking Forum), the physical layer of this protocol is available as a draft standard (ANSI NC ITS T11.1 PH draft 2.6, dated December 2000, ISO/IEC reference number 11518-10) or online at www.hippi.org
This link provides a two-way, 12-channel-wide parallel interface running at 6400 Mbitfs (an increase from the standard HIPPI link rate of 800 Mbit/s).
The link layer uses a fixed size 32-byte packet, 4B/5B encoding, and credit-based flow control, while the physical layer options include parallel copper (to 40 meters) or parallel optics (several hundred meters to 1 km). Relevant standards documents include the following:

ANSI X3.183-mechanical, electrical, and signaling protocol (pH)
ANSI X3.210-framing protocol (FP)
ANSI X3.218-encapsulation of IS0 8802-2 (IEEE 802.2) logical link protocol (LE)
ANSI X3.222-physical switch control (SC)
Serial HIPPI has not been sanctioned as a standard, although various products are available.
HIPPI 6400 (FC-PH)

Synchronous Optical Network (SONET): originally proposed by Bellcore and later standardized by ANSI and ITU (formerly CCITIyr) as ITU-T recommendations G.707, G.708, and G.709.

Related standards include Asynchronous Transfer Mode (ATM), which is controlled by the ATM Forum.

SONET/SDH

The "fundamental rate" of 64 kb/s derives from taking a 4-kHz voice signal (telecom), sampling into 8-bit wide bytes **(**32 kb/s), and doubling to allow for a full-duplex channel (64 kb/s). In other words, this is the minimum data rate required to reproduce a two-way voice conversation over a telephone line. All of the subsequent data rates are standardized as multiples of this basic rate.

DS0	64 kb/s	
T1 = DS1	1.544 Mb/s	24 × DS0
DS1C	3.152 Mb/s	48 × DS0
T2 = DS2	6.312 Mb/s	96 × DS0
T3 = DS3	44.736 Mb/s	672 × DS0
DS4	274.176 Mb/s	4032 × DS0

Note: framing bit overhead accounts for the bit rates not being exact multiples of DSO.

STS/OC is the SONET physical layer ANSI standard. STS refers to the electrical signals and *OC* (optical carrier) to the optical equivalents. Synchronous digital hierarchy (SDH) is the worldwide standard defined by CCITT (now known as ITU, the International Telecommunications Union); formerly known as synchronous transport mode (STM). Sometimes the notation STS-XC is used, where X is the number (1,3, etc.), and ***C*** denotes the frame is concatinated from smaller frames. For example, three STS-1 frames at 51.84 Mb/s each can be combined to form one STS-3C frame at 155.52 Mb/s. Outside the United States, SDH may be called plesiochronous digital hierarchy (PDH). Note that OC and STS channels are normalized to a data rate of 5 1.84 Mbit/s, while the equivalent SDH specifications are normalized to 155.52 Mbit/s.

STS-1 and OC-1	51.840 Mb/s	
STS-3 and OC-3	155.52 Mb/s	same as STM-1
STS-9 and OC-9	466.56 Mb/s	
STS-12 and OC-12	622.08 Mb/s	same as STM-4
STS-18 and OC-18	933.12 Mb/s	
STS-24 and OC-24	1244.16 Mb/s	same as STM-8
STS-36 and OC-36	1866.24 Mb/s	
STS-48 and OC-48	2488.32 Mb/s	same as STM-16
STS-192 and OC-192	9953.28 Mb/s	same as STM-64
STS-256 and OC-256	13271.04 Mb/s	same as STM-86
STS-768 and OC-768	39813.12 Mb/s	same as STM-256
STS-3072 and OC-3072	159252.48 Mb/s	same as STM-1024
STS-12288 and OC-12288	639009.92 Mb/s	same as STM-4096

Higher speed services aggregate low-speed channels by time-division multiplexing; for example, OC-192 can be implemented as four OC-48 data streams. Note that although STS (synchronous transport signal) is analogous to STM (synchronous transport mode), there are some important differences. The first recognized STM is STM-1, which is eqivalent to STS-3. Similarly, not all STS rates have a corresponding STM rate. The frames for STS and STM are both set up in a matrix (270 columns of 9 bytes each), but in STM frames the regenerator section overhead (RSOH) is located in the first 9 bytes of the top 3 rows. The fourth STM row of 9 bytes is occupied by the administrative unit (AU) pointer, which operates in a manner similar to the H1 and H2 bytes of the SONET line overhead (LOH). The 9 bytes of STM frame in rows 5 through 9 is the multiplex section overhead (MSOH) and is similar to the SONET LOH. In SONET we defined virtual tributaries (VT), while SDH defines virtual containers (VC), but they basically work the same way. A VT or VC holds individual El or other circuit data. VCs are contained in tributary unit groups (TUG) instead of SONET'S VT groups (VTG). VCs are defined as follows:

DS-2	VC-2
E3/DS-3	VC-3
DS-1	VC-11
E1	VC-12

Also note that while there are many references to 10 Gbit/s networking in new standards, the exact data rates may vary. As this book goes to press, proposed standards for 10 Gigabit Ethernet and Fibre Channel are not yet finalized. The

Ethernet standard has proposed two data rates, approximately 9.953 Gbit/s (compatible with OC-192) and 10.3125 Gbit/s. Fibre Channel has proposed a data rate of approximately 10.7 Gbit/s. There is also ongoing discussion regarding standards compatible with 40 Gbit/s data rates (OC-768).

The approach used in Europe and elsewhere:

E0	64 Kb/s	
E1	2.048 Mb/s	
E2	8.448 Mb/s	4 E1s
E3	34.364 Mb/s	16 E1s
E4	139.264 Mb/s	64 E1s

Thus, we have the following equivalences for SONET and SDH hierarchies:

Sonet Signal	Sonet Capacity	SDH Signal	SDH Capacity
STS-1, OC-1	28 DS1s or 1 DS3	STM-0	21 E1s
STS-3, OC-3	84 DS1s or 3 DS3s	STM-1	63 E1s or 1 E4
STS-12, OC-12	336 DS1s or 12 DS3s	STM-4	252 E1s or 4 E4s
STS-48, OC-48	1344 DS1s or 48 DS3s	STM-16	1008 E1s or 16 E4s
STS-192, OC-192	5376 DS1s or 192 DS3s	STM-64	4032 E1s or 64 E4s
STS-768, OC-768	21504 DS1s or 768 DS3s	STM-256	16128 E1s or 256 E4s

For completeness, the SONET interface classifications for different applications are summarized in the following table. Recently, a new 300-meter Very Short Reach (VSR) interface based on parallel optics for OC-192 data rates has been defined as well (see Optical Internetworking Forum document OIF2000.044.4, or contact the OIF for details).

	Short Reach	Intermediate Reach	Long Reach		Very Long Reach				
Distance (km)	<2	15	15	40	60	60	120	160	160
Wavelength (nm)	1310	1310	1550	1310	1550	1550	1310	1550	1550
OC-1	SR	IR-1	IR-2	LR-1	LR-2	LR-3	VR-1	VR-2	VR-3
OC-3	SR	IR-1	IR-2	LR-1	LR-2	LR-3	VR-1	VR-2	VR-3
OC-12	SR	IR-1	IR-2	LR-1	LR-2	LR-3	VR-1	VR-2	VR-3
OC-48	SR	IR-1	IR-2	LR-1	LR-2	LR-3	VR-1	VR-2	VR-3
OC-192	SR	IR-1	IR-2	LR-1	LR-2	LR-3	VR-1	VR-2	VR-3

ETHERNET

The major IEEE standards related to Ethernet are listed in the following table. The IEEE defines a common path control (802.1) and datalink layer (802.2) for Ethernet as well as other LAN standards. Although FDDI is controlled by ANSI, it is also intended to fall under the logical link control of IEEE 802.2 (the relevant ISO standard is 8802-2).

Ethernet standards (ref wikipedia, http://en.wikipedia.org/wiki/IEEE_802.3):

Ethernet Standard	Date	Description
Experimental Ethernet	1972	The original Ethernet standard patented in 1978; 2.94 Mbit/s (367 kB/s) over coaxial cable (coax) cable bus
Ethernet II (DIX v2.0)	1982	10 Mbit/s (1.25 MB/s) over thin coax (thinnet)—Frames have a Type field. This frame format is used on all forms of Ethernet by protocols in the Internet protocol suite.
IEEE 802.3	1983	10Base5 10 Mbit/s (1.25MB/s) over thick coax—same as DIX except Type field is replaced by Length, and an 802.2 LLC header follows the 802.3 header.
802.3a	1985	10BASE2 10 Mbit/s (1.25 MB/s) over thin Coax (thinnet or cheapernet)
802.3b	1985	10BROAD36
802.3c	1985	10 Mbit/s (1.25 MB/s) repeater specs
802.3d	1987	FOIRL (Fiber-Optic Inter-Repeater Link)
802.3e	1987	1BASE5 or StarLAN
802.3i	1990	10Base-T 10 Mbit/s (1.25 MB/s) over twisted pair
802.3j	1993	10BASE-F 10 Mbit/s (1.25 MB/s) over Fiber-Optic
802.3u	1995	100BASE-TX, 100BASE-T4, 100BASE-FX Fast Ethernet at 100 Mbit/s (12.5 MB/s) with autonegotiation
802.3x	1997	Full Duplex and flow control; also incorporates DIX framing, so there's no longer a DIX/802.3 split
802.3y	1998	100BASE-T2 100 Mbit/s (12.5 MB/s) over low quality twisted pair
802.3z	1998	1000BASE-X Gbit/s Ethernet over Fiber-Optic at 1 Gbit/s (125 MB/s)
802.3–1998	1998	A revision of base standard incorporating the above amendments and errata
802.3ab	1999	1000BASE-T Gbit/s Ethernet over twisted pair at 1 Gbit/s (125 MB/s)
802.3ac	1998	Max frame size extended to 1522 bytes (to allow "Q-tag") The Q-tag includes 802.1Q VLAN information and 802.1p priority information.
802.3ad	2000	Link aggregation for parallel links
802.3-2002	2002	A revision of base standard incorporating the three prior amendments and errata
802.3ae	2003	10 Gbit/s (1250 MB/s) Ethernet over fiber; 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-SW, 10GBASE-LW, 10GBASE-EW
802.3af	2003	Power over Ethernet
802.3ah	2004	Ethernet in the First Mile

Continued

Ethernet Standard	Date	Description
802.3ak	2004	10GBASE-CX4 10 Gbit/s (1250 MB/s) Ethernet over twin-axial cable
802.3–2005	2005	A revision of base standard incorporating the four prior amendments and errata.
802.3an	2006	10GBASE-T 10 Gbit/s (1250 MB/s) Ethernet over unshielded twisted pair(UTP)
802.3ap	exp. 2007	Backplane Ethernet (1 and 10 Gbit/s (125 and 1250 MB/s) over printed circuit boards)
802.3aq	2006	10GBASE-LRM 10 Gbit/s (1250 MB/s) Ethernet over multimode fiber
802.3ar	exp. 2007	Congestion management
802.3as	2006	Frame expansion
802.3at	exp. 2008	Power over Ethernet enhancements
802.3au	2006	Isolation requirements for Power Over Ethernet (802.3-2005/Cor 1)
802.3av	exp. 2009	10 Gbit/s EPON
802.3 HSSG	exp. 2009	Higher Speed Study Group. 100 Gb/s up to 100 m or 10 km using MMF or SMF optical fiber, respectively

Fast Ethernet (10 Mbit/s).

100Base-T	Copper	100 m	IEEE 802.3
100Base-TX	Copper, twin pair cat 5 or better		
100Base-T4	Copper, 4 pair cat 3 or better		
100Base-T2	Copper, 2 pair cat 3 or better		
100Base-FX	Multimode, 1300 nm (LED source)	400 m half duplex, 2 km full duplex	Not backward compatible with 10Base-FL
100Base-SX	Multimode, 850 nm	300 m	Backward compatible with 10Base-FL
100Base-BX	Single-mode, over 1 fiber with 2 wavelength CWDM		

Gigabit Ethernet.

1000Base-SX	Multimode, 850 nm	220 m over 62.5 micron fiber; typically 500 m over 50 micron fiber	IEEE 802.3z Max TX = −5 dBm Min RX = −14 dBm
1000Base-LX	Single-mode, 1300 nm	2 km	
	Multimode, 1300 nm	550 m over 50 micron fiber (requires mode conditioners over 300 m)	
1000Base-CX	Copper	25 m	Balanced copper shielded twisted pair version; nearly obsolete
1000Base-ZX or 1000Base-LH	Single-mode, 1550 nm	70 km	nonstandard
1000Base-T	Copper	100 m over cat 5 or higher	IEEE 802.3ab
1000Base-TX	Copper	Same as 1000Base-T	TIA standard, often confused with 1000Base-T

10 Gigabit Ethernet Physical Media Dependent Sublayers (PMDs).

10GBase-E	Single-mode, 1550 nm	Up to 40 km	
10GBase-L	Single-mode, 1300 nm	Up to 10 km	
10GBase-S	Multimode, 850 nm	26–82 m on OM2 fiber 300 m on OM3 fiber	
10Gbase-LX4	Single-mode, 4 wavelength CWDM	10 km	4 transmitters × 3.125 Gbit/s each
	Multimode, 4 wavelength CWDM	300 m	4 transmitters × 3.125 Gbit/s each

10 Gigabit Ethernet LAN PHY (uses 64/66B encoding, line rate of 10.3 Gbit/s.

10GBase-SR	Multimode, 850 nm	26–82 m on OM2 fiber 300 m on OM3 fiber	IEEE 802.3ae
10GBase-LRM	Multimode, 850 nm	220 m over FDDI grade (62.5 micron) fiber	IEEE 802.3aq
10GBase-LR	Single-mode, 1300 nm	10 km	IEEE 802.3 clause 48 (8B/10B 4 channel parallel bridge) for Xenpak, X2, Xpak; IEEE802.3 clause 49 (64/66B serial) for XFP
10GBase-ER	Single-mode, 1550 nm	40 km	
10GBase-ZR	Single-mode, 1550 nm	80 km	Not specified by IEEE 802; based on OC-192/STM-64 SONET/SDH specs
10GBase-LX4	Multimode, CWDM near 1300 nm	240–300 m	4 transmitters × 3.125 Gbit/s each
	Single-mode, CWDM near 1300 nm	10 km	4 transmitters × 3.125 Gbit/s each

10 G ETHERNET WAN PHY

10GBASE-SW, 10GBASE-LW and 10GBASE-EW are varieties that use the WAN PHY, designed to interoperate with OC-192/STM-64 SDH/SONET equipment using a lightweight SDH/SONET frame running at 9.953 Gbit/s. WAN PHY is used when an enterprise user wishes to transport 10 G Ethernet across telco SDH/SONET or previously installed wave-division multiplexing systems without having to directly map the Ethernet frames into SDH/SONET. The WAN PHY variants correspond at the physical layer to 10GBASE-SR, 10GBASE-LR, and 10GBASE-ER, respectively, and hence use the same types of fiber and support the same distances. There is no WAN PHY standard corresponding to 10GBASE-

LX4 and 10GBASE-CX4 since the original SONET/SDH standard requires a serial implementation.

10 Gigabit Ethernet Copper Interfaces.

10GBase-CX4	4 lanes parallel	15 m	IEEE 802.3ak
10GBase-KX	-KR (same coding as 10GBase-LR/ER/SR with optional FEC -KX4 (same as 10GBase-CX4)	40 inches copper PCB with 2 connectors	Backplane Ethernet, IEEE 802.3ap
10GBase-T		55 m over cat 6 (proposed 100 m over cat 6a)	IEEE 802.3an

ETHERNET FIRST MILE STANDARDS

Ethernet First Mile over Copper (EFMC)—cat 3 copper, 705 m at 10 Mbit/s or 2.7 km at 2 Mbit/s

Ethernet First Mile over Fiber (EFMF)—single-mode fiber, 10 km, 100 Mbit/s or 1000 Mbit/s

EFM passive optical networks (EFMP)—single-mode fiber point-to-multipoint networks, 20 km, 1000 Mbit/s

EFMF Standards—Dual Fiber (2 fibers used for transmit and receive).

100Base-LX10	Single-mode, 1300 nm	10 km, 125 Mbd (transceivers compatible with OC-3/STM-1)	Tx = –15 dBm RX = –25 dBm Uses 4B/5B coding, NRZI
1000Base-LX10	Single-mode, 1300 nm	10 km, 1250 Mbd	TX = –9.5 dBm RX = –20 dBm Uses 8B/10B coding

EFMF Standards—Single Fiber (bidirectional transmission with 2 wavelengths on one fiber, 1300 nm uplink, 1500 nm downlink; uplink denoted by -U, downlink by -D).

100Base-BX10-U	Single-mode, 1300 nm	10 km, 125 Mbd	TX = –14 dBm RX = –29.2 dBm
100Base-BX10-D	Single-mode, 1500 nm		
1000Base-BX10-U	Single-mode, 1300 nm	10 km, 1250 Mbd	TX = –9 dBm RX = –20 dBm
1000Base-BX10-D	Single-mode, 1500 nm		

Metro Ethernet Forum (MEF) Carrier Class Ethernet Release Levels

- MEF 2 Requirements and Framework for Ethernet Service Protection
- MEF 3 Circuit Emulation Service Definitions, Framework and Requirements in Metro Ethernet Networks
- MEF 4 Metro Ethernet Network Architecture Framework

Part 1: Generic Framework

- MEF 6 Metro Ethernet Services Definitions Phase I
- MEF 7 EMS-NMS Information Model
- MEF 8 Implementation Agreement for the Emulation of PDH Circuits over Metro Ethernet Networks
- MEF 9 Abstract Test Suite for Ethernet Services at the UNI
- MEF 10.1 Ethernet Services Attributes Phase 2*
- MEF 11 User Network Interface (UNI) Requirements and Framework
- MEF 12 Metro Ethernet Network Architecture Framework

Part 2: Ethernet Services Layer

- MEF 13 User Network Interface (UNI) Type 1 Implementation Agreement
- MEF 14 Abstract Test Suite for Traffic Management Phase 1
- MEF 15 Requirements for Management of Metro Ethernet

Phase 1 Network Elements

- MEF 16 Ethernet Local Management Interface
- MEF 10.1 replaces and enhances MEF 10 Ethernet Services Definition Phase 1 and replaced MEF 1 and MEF 5.

There are several complementary industry standards used by the MEF:

- IEEE provider bridge 802.1ad and provider backbone bridge 802.1ah
- ITU-T SG15 references Ethernet private line and Ethernet virtual private line specifications
- IETR layer 2 VPNs
- IEEE 802.1ag fault management and 802.3ah link OAM
- ITU-T SG13 (service OAM), harmonized with SG4
- OIF customer signaling of Ethernet services
- IETF MPLS fast reroute, graceful restart

Carrier Ethernet Attributes.

MEF	Standardized Service	Service Management	Reliability	Quality of Service	Scalability
MEF2			Architecture Area		
MEF3	Service Area			Service Area	
MEF4	Architecture Area				
MEF6	Service Area			Service Area	Service Area
MEF7		Management Area			
MEF8	Service Area				
MEF9	Test & Measurement		Test & Measurement		
MEF10	Service Area			Service Area	Service Area
MEF11	Architecture Area				
MEF12	Architecture Area				Architecture Area
MEF13	Architecture Area				
MEF14	Test & Measurement		Test & Measurement	Test & Measurement	
MEF15		Management Area			
MEF16		Management Area			

Common Fiber-Optic Attachment Options (without repeaters or channel extenders).

Channel	Fiber	Connector	Bit Rate	Fiber Bandwidth	Maximum Distance	Link Loss
ESCON (SBCON)	SM	SC duplex	200 Mb/s	N/A	20 km	14 dB
	MM 62.5 micron	ESCON duplex or MT-RJ	200 Mb/s	500 MHz-km	2 km	8 dB
				800 MHz-km	3 km	
	MM 50.0 micron	ESCON duplex or MT-RJ	200 Mb/s	800 MHz-km	2 km	8 dB
Sysplex Timer ETR/CLO	MM 62.5 micron	ESCON duplex or MTRJ	8 Mb/s	500 MHz-km or more	3 km	8 dB
	MM 50.0 micron	ESCON duplex or MTRJ	8 Mb/s	500 MHz-km or more	2 km	8 dB
FICON/ Fibre Channel LX	SM	SC duplex or LC duplex	1.06 Gb/s	N/A	10 km	7 dB

Continued

Common Fiber-Optic Attachment Options (without repeaters or channel extenders) (*Continued*)

Channel	Fiber	Connector	Bit Rate	Fiber Bandwidth	Maximum Distance	Link Loss
FICON LX	MM w/MCP 62.5 micron	SC duplex or LC duplex	1.06 Gb/s	500 MHz-km	550 meters	5 dB
	MM w/MCP 50.0 micron	SC duplex or LC duplex	1.06 Gb/s	400 MHz-km	550 meters	5 dB
FICON/ Fibre Channel SX	MM 50.0* micron	SC duplex or LC duplex	1.06 Gb/s	500 MHz-km	500 meters	3.85 dB
	MM 62.5* micron	SC duplex or LC duplex	1.06 Gb/s	160 MHz-km	250 meters	2.76 dB
	MM 62.5* micron	SC duplex or LC duplex	1.06 Gb/s	200 MHz-km	300 meters	3 dB
	MM 50.0* micron	SC duplex or LC duplex	2.1 Gb/s	500 MHz-km	300 meters	
	MM 62.5* micron	SC duplex or LC duplex	2.1 Gb/s	160 MHz-km	120 meters	
	MM 62.5* micron	SC duplex or LC duplex	2.1 Gb/s	200 MHz-km	150 meters	
Parallel Sysplex coupling links- HiPerlinks	SM	SC duplex or LC duplex	1.06 Gb/s compatibility mode	N/A	10 km (20 km on special request)	7 dB
	SM	LC duplex	2.1 Gb/s peer mode	N/A	10 km (20 km on special request)	7 dB
	MM w/MCP 50.0 micron	SC duplex or LC duplex	1.06 Gb/s	500 MHz-km	550 meters	5 dB
	MM 50.0* Discontinued May 98	SC duplex	531 Mb/s	500 MHz-km*	1 km	8 dB*
OC-3/ATM 155	SM	SC duplex	155 Mb/s	N/A	20 km	15 dB
	MM 50 micron	SC duplex	155 Mb/s	500 MHz-km	2 km	
	MM 62.5 micron	SC duplex	155 Mb/s	500 MHz-km	2 Km	11 dB

Continued

Common Fiber-Optic Attachment Options (without repeaters or channel extenders) (*Continued*)

Channel	Fiber	Connector	Bit Rate	Fiber Bandwidth	Maximum Distance	Link Loss
	MM* 50 micron	SC duplex	155 Mb/s	500 MHz-km	1 km	
	MM* 62.5 micron	SC duplex	155 Mb/s	160 MHz-km	1 km	
OC-12/ATM 622	MM 50 micron	SC duplex	622 Mb/s	500 MHz-km	500 m	
	MM 62.5 micron	SC duplex	622 Mb/s	500 MHz-km	500 m	
	MM* 50 micron	SC duplex	622 Mb/s	500 MHz-km	300 m	
	MM* 62.5 micron	SC duplex	622 Mb/s	160 MHz-km	300 m	
FDDI	MM 62.5 micron	Media Access Connector (MAC) or SC duplex	125 Mbit/s overhead reduces to 100 Mb/s	500 MHz-km	2 Km	9 dB
	MM 50 micron	MAC or SC duplex	125 Mbit/s overhead reduces to 100 Mb/s	500 MHz-km	2 km	9 dB
Token Ring*	MM 62.5 micron	SC duplex	16 Mbit/s	160 MHz-km	2 km	
	MM 50 micron	SC duplex	16 Mbit/s	500 MHz-km	1 km	
Ethernet	MM* 50 micron 10Base-F	SC duplex	10 Mbit/s	500 MHz-km	1 km	
	MM* 62.5 micron 10Base-F	SC duplex	10 Mbit/s	160 MHz-km	2 km	
	MM* 50 micron 100Base-SX	SC duplex	100 Mbit/s	500 MHz-km	300 m	
	MM* 62.5 micron 100Base-SX	SC duplex	100 Mbit/s	160 MHz-km	300 m	
Fast Ethernet	MM 50 micron 100Base-F	SC duplex	100 Mbit/s	500 MHz-km	2 km	
	MM 62.5 micron 100Base-F	SC duplex	100 Mbit/s	500 MHz-km	2 km	

Continued

Common Fiber-Optic Attachment Options (without repeaters or channel extenders) (*Continued*)

Channel	Fiber	Connector	Bit Rate	Fiber Bandwidth	Maximum Distance	Link Loss
Gigabit Ethernet IEEE 802.3z	SM 1000BaseLX	SC duplex	1.25 Gb/s	N/A	5 km	4.6 dB
	MM* 62.5 micron 1000BaseSX	SC duplex	1.25 Gb/s	160 MHz-km 200 MHz-km	220 meters 275 meters	2.6 dB *
	MM w/MCP 62.5 micron 1000BaseLX	SC duplex	1.25 Gb/s	500 MHz-km	550 meters	2.4 dB
	MM* 50.0 micron 1000BaseSX	SC duplex	1.25 Gb/s	500 MHz-km*	550 meters	3.6 dB*
	MM w/MCP 50.0 micron 1000BaseLX	SC duplex	1.25 Gb/s	500 MHz-km	550 meters	2.4 dB

Notes:

*Indicates channels that use short-wavelength (850 nm) optics; all link budgets and fiber bandwidths should be measured at this wavelength.

*SBCON is the non-IBM trademarked name of the ANSI industry standard for ESCON.

*All industry standard links (ESCON/SBCON, ATM, FDDI, Gigabit Ethernet) follow published industry standards. Minimum fiber bandwidth requirement to achieve the distances listed is applicable for multimode (MM) fiber only. There is no minimum bandwidth requirement for single mode (SM) fiber.

*Bit rates given below may not correspond to effective channel data rate in a given application due to protocol overheads and other factors.

*SC duplex connectors are keyed per the ANSI Fiber Channel Standard specifications.

*MCP denotes mode conditioning patch cable, which is required to operate some links over MM fiber.

*As light signals traverse a fiber-optic cable, the signal loses some of its strength (decibels (dB) is the metric used to measure light loss). The significant factors that contribute to light loss are: the length of the fiber, the number of splices, and the number of connections. All links are rated for a maximum light loss budget (i.e., the sum of the applicable light loss budget factors must be less than the maximum light loss budget) and a maximum distance (i.e., exceeding the maximum distance will cause undetectable data integrity exposures). Another factor that limits distance is jitter, but this is typically not a problem at these distances.

*Unless noted, all links are long wavelength (1300 nm), and the link loss budgets and fiber bandwidths should be measured at this wavelength. For planning purposes, the following worst case values can be used to estimate the total fiber link loss. Contact the fiber vendor or use measured values when available for a particular link configuration:

Link loss at 1300 nm = 0.50 dB/km
Link loss per splice = 0.15 dB/splice (not dependent on wavelength)
Link loss per connection = 0.50 dB/connection (not dependent on wavelength)

*HiPerLinks are also known as Coupling Facility (CF) or Inter System Channel (ISC) links.

*All links may be extended using channel extenders, repeaters, or wavelength multiplexers. Wavelength multiplexing links typically measure link loss at 1500 nm wavelength, typical loss is 0.3 dB/km.

Transmit and Receive Levels of Common Fiber-Optic Protocols.

Protocol type	I/O Spec.
ESCON/SBCON MM and ETR/CLO MM	TX: –15 to –20.5
	RX: –14 to –29
EXCON/SBCON SM	TX: –3 to –8
	RX: –3 to –8
FICON LX SM (MM via MCP)	TX: –4 to –8.5
	RX: –3 to –22
FICON SX MM	TX: –4 to –9.5
	RX: –3 to –17
ATM 155 MM (OC-3)	TX: –14 to –19
	RX: –14 to –30
ATM 155 MM (OC-3)	TX: –8 to –15
	RX: –8 to –32.5
FDDI MM	TX: –14 to –19
	RX: –14 to –31.8
Gigabit Ethernet LX SM (MM via MCP)	TX: –14 to –20
	RX: –17 to –31
Gigabit Ethernet SX MM (850 nm)	TX: –4 to –10
	RX: –17 to –31
HiPerLinks (ISC coupling links, IBM Parallel Sysplex, 1.06 Gbit/s peer mode)	TX: –3 to –11
	RX: –3 to –20
HiPerLinks (ISC coupling links, IBM Parallel Sysplex, 2.1 Gbit/s peer mode)	TX: –3 to –9
	RX: –3 to –20

This page intentionally left blank

Appendix E

Other Datacom Developments

Because the field of fiber-optic data communication is expanding so rapidly, inevitably new technologies and applications will emerge while this book is being prepared for print. In addition, there will be some recent developments that have not been incorporated into the previous chapters. This appendix is an attempt to address these changes by including a partial list of related datacom devices and standards for reference purposes and presenting brief descriptions of several emerging datacom technologies that the reader may encounter (product names and terminology used in this appendix may be copyrighted by the companies that developed them).

FREE-SPACE OPTICAL LINKS

Recent advances have made free-space optical data links practical in conditions when the weather is not a factor in attenuating the signal. Although bit error rates are typically on the order of 10e-6 to 10e-9 at 155 MB/s, this is adequate for some applications such as voice and video transmission. Some experiments have shown that error rates as low as 10e-l2 can be obtained at 200 MB/s using ESCON protocols on free-space optical links. For example, AstroTerra Corporation currently offers the TerraLink system, a tripod-mounted line-of-sight communication link with an 8-in. telescope aperture capable of 155 MB/s transmission over 8 km in clear weather. Other companies have demonstrated 1.2 Gbit/s freespace communication over 150 km. This technology may form the basis of a communication system between satellites or aircraft in flight. Portable free-space laser communications binocular communications systems at 100 kb/s over 3–5 km are sometimes used as part of disaster recovery efforts following hurricanes or earthquakes.

OPTICAL CONNECTORS

Older connector types, including the optical SMA connector, DIN, and Light-Ray MPX (a 2- to 12-element multifiber connector incompatible with the MTP/MPO) are far less common than they were a decade ago. However, they may still be found occasionally in some installations. Many other legacy connectors continue to be used in reasonable quantities, including the SC and ST connectors (recall that the ST was developed and trademarked by AT&T over 20 years ago). As new connector options emerge, many types of parallel or multifiber connectors have been proposed, not all of which are intermateable. For example, the *MTP* connector and its variations (*MPO* or MPX) do not mate with the SMC connector used on the Paroli parallel optical transceivers, yet both are being deployed as part of HIPPI 6400 installations.

Optical backplane connectors are likewise not yet standardized and may include variants such as the MAC-II (an improved version of the original 12-fiber AT&T MAC connector) or the mini-MAC (a Bellcore approved version based on the MT ferrule). There have been various proposals for stacking multiple MT fermles in a single connector; for example, a stack of 6 MTs forming a 72-fiber connector has been developed, and multichannel optical transceivers capable of utilizing this capacity in a full duplex 320 Gbit/s link have been demonstrated. Research efforts continue to push the limits of stacked one-dimensional and two-dimensional connectors; an example is the 144-fiber "super-MT," which has been proposed, and the 4×12 element Diamond MF multiple fiber connector. As optical interconnects are designed into computer packages, right-angled connectors are being investigated; these are expected to play a role in future supercomputers and mainframes.

TRANSCEIVERS

Historically, there has always been interest in pluggable electrical and/or optical interfaces. Some of the early examples included the gigabit interface connector (GBIC), which can still be found in some systems today. This was a gigabit module that offerred the convenience of being able to unplug a copper interface and replace it with an optical interface at a later time for increased distance and bandwidth. It featured a parallel electrical interface; the copper and optical modules plug into the same electrical connection point, which includes a pair of rails to guide the modules during plugging. The name GBIC has sometimes been used in reference to any pluggable transceiver interface, even the more recent SFP form factors.

Low-cost bidirectional transceivers for fiber in the loop and similar applications (fiber to the curb or home) have also been developed. For example, one such device offers a connectorized, single-mode bidirectional transceiver with a source, modulator, and detector integrated into a single package. Each module includes

an InP diode laser, silicon CMOS modulator/demodulator electronics, a PIN photodiode, an integrated beam splitter, and a built-in SC for ferrule-based fiber attachment. The beam splitter is used for bidirectional transmission at 1.55-pm wavelength and reception at 1.3-pm wavelength. Data rates up to 1.2 Gb/s are 1.3 pm and a burst mode, 50 MB/s receiver at 1.55 pm, are available in a package measuring $85 \times 17 \times 10$ mm.

There are a number of trends in the future development of datacom transceiver technology. Among these is the migration to lower power supply voltages, following the trend of digital logic circuits. Most digital logic has migrated from 5 to 3.3 V and is on a path toward 2.5 V operation. Many optical transceivers can operate at either 3.3 V or 5 V, though some require both voltages; lower voltage devices are still under development.

Although the telecommunications market continues to have distinct requirements from the datacommunications market, there is certainly some merging of requirements between these two areas. The telecom market is currently driving development of analog optical links for cable television applications, including bidirectional wavelength multiplexed links that transmit in one direction at 1300 nm and in the other at 1550 nm. This is achieved by incorporating some form of optical beam splitter in the transceiver package. Many of these devices incorporate optical fiber pigtails rather than connectorized transceiver assemblies. Packaging is another evolving area; existing datacom industry standards, such as the 1×9 pin serial transceiver or 2×9 pin transceiver with integrated clock recovery, have given way to surface mount cages for pluggable transceivers. In the telecom industry, the standard optical transceiver package has been the dual inline pin (DIP) package (14 pins and two rows); recent developments in the so-called mini-DIP (8 pins and two rows) represent another packaging variant. Finally, integrated optoelectronic integrated circuits are advancing to the point where they may replace the standard TO can and ball lens assembly for optical device packaging and alignment. VCSEL**s** continue to hold a great deal of promise as low-cost optical sources for data communications, especially for optical array interconnections. Research continues on VCSELs that can operate at long wavelengths (1300 nm).

This page intentionally left blank

Appendix F

Laser Safety

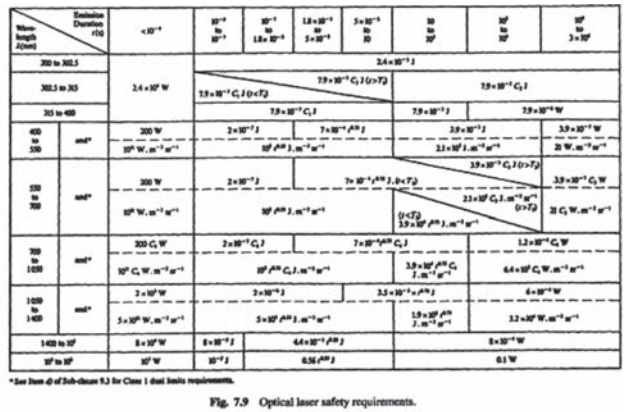
Laser safety is an important criterion for many data communication products because they are accessible to customers who may have little or no laser training. This limits the amount of optical power that can safely be launched into the link and may impact distance and BER performance. This is a very important area; infrared sources of less than 1 mW can cause serious damage to the unprotected eye. There are two major laser eye safety standards:

1. The international standard IEC 60825- 1 of the International Electrotechnical Commission, applicable if a laser product or component is to be distributed worldwide.
2. The U.S. national regulation 21 CFR, Chapter 1, Subchapter J, of the Center for Devices and Radiological Health (CDRH), a subgroup of the Department of Health and Human Services of the Occupational Safety and Health Administration and the U.S. Food and Drug Administration (FDA), if the laser product is for the U.S. market. This standard is also published by the American National Standards Institute (ANSI).

In some cases, there may be additional local or state labor safety regulations as well. For example, the New York State Department of Labor enforces New York Code Rule 50, which requires all manufacturers to track their primary laser components by a state-issued serial number. In general, the FDA and IEC standards define acceptable laser eye safety classifications for different

types of lasers at different operating wavelengths. Historically, in the United States, the base standard, ANSI 2136.1 (“Standard for the safe use of lasers”) was first released in 1988 and has been revised periodically since then. The related standard, ANSI 2136.2 (“Standard for the safe use of optical fiber communication systems utilizing laser diode and LED sources”), has also been updated to reflect the changes to 2136.1. A related document, ANSI 2136.3 (“Standard for the safe use of lasers in health care facilities”), was reissued in April 1996 with some new applications and modest changes to control procedures. Three related ANSI standards are 2136.4, which covers the measurement of laser power and energy; 2136.5, which covers safety aspects of lasers used in educational institutions; and 2136.6, which covers outdoor use of lasers for applications such as surveying, laser displays, and military systems.

Ideally, datacom applications require that all fiber-optic products be class I, or inherently safe; there is no access to unsafe light levels during normal operation or maintenance of the product or during accidental viewing of the optical source. This includes viewing the optical fiber end face while the transmitter is in operation. In the United States, the FDA standard for a class 1 product is limited to −2.0 dBm output power. International standards defined by the IEC are somewhat different and limit class 1 exposure to −6.0 dBm. Following recent revisions of the laser safety standards, the FDA requirements currently permit lower optical power levels and higher power densities than the IEC standards; these differences are illustrated in the accompanying figure. Note that the IEC requires that a product must operate as a class 1 device even under a single point of failure. This requires redundancy in the hardware design. For example, some early transceivers used a 1×9 pinout, which was later modified to a 2×9 pinout to implement clock recovery and redundant laser safety. More recent laser technologies such as as vertical cavity surface-emitting laser (VCSEL) arrays produce a beam of circular cross section rather than the elliptical beam profile typical of long-wavelength lasers. As a result, most of the optical power is launched into the lower order modes of an optical fiber and tends to remain there even after fairly long distances. Although this improves modal noise performance, it also makes laser safety compliance more difficult. Some products treat these as extended sources, or (as we will discuss shortly) they may be eligible for reclassification as class 1M products. A summary of the different laser safety classifications and power levels is given in the accompanying figures; for details and current revisions of all standards, contact the referenced standards bodies, since these standards undergo regular revisions. Recent revisions in the standards apply to all optical sources, including light-emitting diodes (LEDs), both surface-emitting diodes and the edge-emitting diodes required for higher speed applications. Laser safety power limits are given in the following tables:

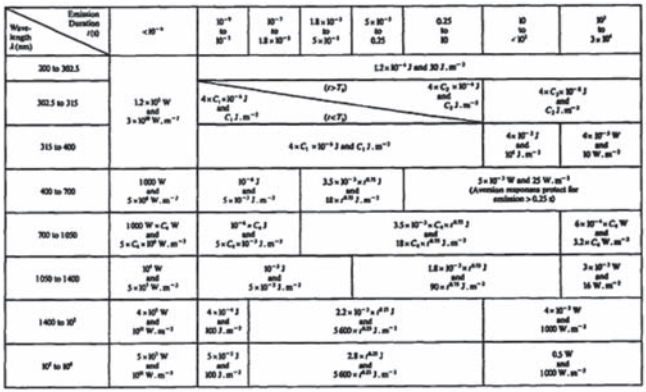


*See Item *d* of Sub-clause 9.3 for Class 1 dual limits requirements.

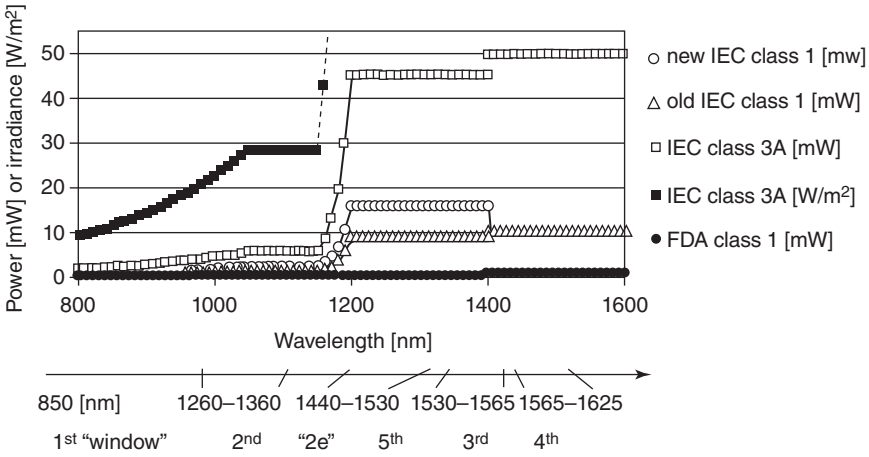
260

Fig. 7.9 Optical laser safety requirements.

Wavelength λ (nm)	Emission duration *t* (s)	Class 2 AEL
400 to 700	$t < 0.25$	Same as Class 1 AEL
	$t < 0.25$	10^{-3} W



Wave-length λ (nm) \ Emission duration t (s)	$<10^{-9}$	10^{-9} to 0.25	0.25 to 3×10^4
200 to 302.5	3.8×10^4 W	3.8×10^{-4} J	1.5×10^{-3} W
302.5 to 315	1.25×10^6 C_2 W	1.25×10^{-3} C_2 J	5×10^{-5} C_2 W
315 to 400	1.25×10^8 W	0.125 J	0.5 W
400 to 700	3.14×10^{11} W.m^{-2}	3.14×10^3 $t^{0.75}$ J.m^{-2} and $<10^5$ J.m^{-2}	0.5 W
700 to 1050	3.14×10^{11} C_2 W.m^{-2}	3.14×10^5 C_2 $t^{0.75}$ J.m^{-2} and $<10^5$ J.m^{-2}	0.5 W
1050 to 1400	1.57×10^{12} W.m^{-2}	1.57×10^6 $t^{0.75}$ J.m^{-2} and $<10^5$ J.m^{-2}	0.5 W
1400 to 10^6	10^{14} W.m^{-2}	10^5 J.m^{-2}	0.5 W



Power/Irradiance limits according to IEC and FDA standards.

ANSI Z136.1 (1993) Laser Safety Classifications.

Class 1—Inherently safe; no viewing hazard during normal use or maintenance; 0.4 μW or less; no controls or label requirements

Class 2—Human aversion response sufficient to protect eye; low-power visible lasers with power less than 1 mW continuous; in pulsed operation, power levels exceed the class 1 acceptable exposure limit for the entire exposure duration but do not exceed the class 1 limit for a 0.25 second exposure; requires caution label

Class 2a—Low-power visible lasers that do not exceed class 1 acceptable exposure limit for 1000 s or less (not intended for viewing the beam); requires caution label

Class 3a—Aversion response sufficient to protect eye unless laser is viewed through collecting optics; 1–5 mW; requires labels and enclosure/interlock; warning sign at room entrance

Class 3b—Intrabeam (direct) viewing is a hazard; specular reflections may be a hazard; 5–500 mW continuous; <10 J per square centimeter pulsed operation (less than 0.25 s) same label and safety requirements as class 3a plus power actuated warning light when laser is in operation

Class 4—Intrabeam (direct) viewing is a hazard; specular and diffuse reflections may be a hazard; skin protection and fire potential may be concerns; >500 mW continuous; >10 J per square centimeter pulsed operation; same label and safety requirements as class 3b plus locked door, door actuated power kill, door actuated filter, door actuated shutter, or equivalent.

Note: Laster training is required in order to work with anything other than class 1 laser products.

Following a restructuring of the safety standards in 2001, class 1, 2, 3b, and 4 lasers remained unchanged, and a new class 1M was introduced that is considered safe under reasonably foreseeable conditions if optical instruments are not used for viewing. There is no upper limit on the total output power of class 1M laser products; they are likely to include lasers or LEDs with highly divergent

beams or lasers with wide-aperture collimated beams. Another recent update, the class 2M laser, operates in the visible range only, and is considered safe if no optical instruments are used for viewing and the blink or aversion response operates. Some current class 3b laser products (capable of causing injury) may qualify for reclassification as either class 1M or 2M; in particular, this may affect parallel optical interconnects. Furthermore, a new class 3R laser is defined as having accessible emissions that exceed the MPE for exposures of 0.25 second if they are visible and for 100 seconds if they are invisible. Total output power for a class 3R laser must not exceed the accessible emission limits for class 2 (visible) or class 1 (invisible) by more than a factor of *5*.

Class 1 laser certification is granted by testing a product at an independent laboratory. In the United States, products are granted an accession number by the FDA to certify their compliance with class 1 limits. It is a popularly held misconception that the FDA issues "certification" for products such as fiber-optic transceivers that contain laser devices. This is not the case. Manufacturers of such equipment provide a report and product description to the FDA, and the manufacturer certifies through supporting test results that the product complies with class 1 regulations; the FDA then assigns an accession number in order to track this part and permit import-export as a class 1 laser device. International certification is usually performed by a recognized testing laboratory, such as TUV or VDE in Germany. These labs will assign approval numbers to the products in question based on their independent testing of samples provided by the manufacturer and assessment of any supporting manufacturer data. A fee may be associated with this evaluation, and an annual fee must be paid in order to keep an IEC registration number active. Otherwise, it is withdrawn, and the product can no longer claim to be IEC class 1 compliant. International safety standards define the requirements for labels that must be affixed to class 1 optical products and the terminology that must be used in product safety literature.

In order to receive class 1 certification worldwide, a product must be single-fault tolerant; that is, the optical power cannot exceed recommended levels if a single point of failure occurs in the link. There are three common ways in which this can be guaranteed by the manufacturer. First, the optical transceiver can be designed to never emit more than the maximum allowed optical power level. This is the simplest approach, but is only possible due to recent relaxation in the laser safety standards; products manufactured prior to 1994 required alternate solutions. Some products use higher power laser transmitters but never couple more than the maximum power level into the optical fiber; this accounts for coupling loss in the transceiver and does not present a hazard for viewing the fiber end face. In this case, the transmitter must have some additional form of protection, such as mechanical shutters that snap into place covering the laser when not in use. This design was used on some single-mode ESCON transceivers prior to 1995. Finally, the ANSI Fibre Channel Standard defines a laser safety method

called Open Fiber Control (OFC). This implements an interlock at the transceiver that detects whether light is being received; both ends of a duplex link must detect light and exchange a handshake sequence before the laser transmitter is activated. If the link is opened for any reason, such as a broken fiber or pulled connector, the lasers turn off on both sides of the link. The link then enters a waiting mode, in which a low-power optical pulse is transmitted every 10 s. When the link is closed once again, this pulse is detected by both ends of the link, which exchange handshakes and activate the laser transmitters at full power again. Using this approach, the link can employ much higher powered lasers without violating class 1 certification limits; this can enable longer distances and improved BER performance. This approach was used on early Fibre Channel products but is no longer widely used.

LASER SAFETY CERTIFICATION INFORMATION

International Electrotechnical Commission (IEC): IEC International
Laser Safety & Compliance Laboratories
VDE (Verband Deutscher Electrotechniker)
Association of German Electrical Engineers
Merianstrabe 28 D-6050 Offenbach
Phone: +49 069 8306-0

TUV Rheinland of North America, Inc. (U.S. contact for TUV Laboratory)
12 Commerce Road
Newton, CT 06470
Phone: (203) 426-0888
TUV Rheinland
Prufstelle fur Geratesicherheit
Am Grauen Stein
5000 Koln 91
Germany

U.S. Food and Drug Administration (FDA)
Center for Devices and Radiological Health
2098 Gaither Road
Rockville, MD 20850

The FDA issues accession numbers to verify the compliance of laser and optical transmitter products with regulations for the administration and enforcement of the Radiation Control for Health and Safety Act of 1968 (Title 21, CFR, Subchapter J); the U.S. laser safety certification is obtained from the standard FDA 21 CFR Part 1040, performance standards for light-emitting products Sections 1040.10 (laser products) and 1040.1 1 (specific-purpose laser products)

Certified Laser Safety Officer Training
Laser Institute of America (LIA)
12424 Research Parkway
Suite 125
Orlando, FL 32826
Phone: (407) 380-1553
(In addition to providing laser safety training, copies of the ANSI Z136.X laser safety standards are available from the LIA).

This page intentionally left blank

Index

Accelerated Strategic Computing Initiative (ASCI), 448
Acousto-optic tunable filter, tunable receiver mechanisms, 639
Active optical cable, 82–83
Alignment
  angular alignment, 200–202
  automated alignment
    active, 208–209
    hybrid active/passive, 210–214
    passive, 214–216
  costs, 206–207
  module alignment, 202
  optical alignment, 197
  planar alignment, 198–200
  requirements, 207
  technology, 207
  variables and interdependence, 203–206
Amplification, detector noise, 157–158
Amplifier, *see* Optical amplifiers
Amplifier spontaneous emission (ASE), 642
Angle Polished Connector (APC), 121
Angular alignment, 200–202
Angular frequency, photocurrent, 145
APC, *see* Angle Polished Connector
APD, *see* Avalanche photodiode
APON, *see* ATM passive optical network
Arrayed waveguide grating (AWG), 374
ASCI, *see* Accelerated Strategic Computing Initiative
ASE, *see* Amplifier spontaneous emission
Asynchronous transfer mode (ATM)
  cell versus circuit switching, 484–486
  cell versus packet switching, 482, 484
  classical IP over ATM implementation, 492–493
  layer
    overview, 490–491
    virtual channels and paths, 491–492
  layered architecture, 486–487
  local area network emulation
    components, 494
    definition, 493
    LAN emulation client
      control connections, 495
      data connections, 495–496
      initialization phases, 494
    objectives, 493
    operations, 496–498
  overview, 475
  physical layers
    cell stream, 489
    media requirements, 489
    overview, 487–488
    payload capacity comparison versus SONET, 490
    SONET/SDH, 487–489
ATM, *see* Asynchronous transfer mode
ATM address resolution protocol (ATMARP), 492–493
ATM passive optical network (APON), 413–414
ATMARP, *see* ATM address resolution protocol
Attenuation
  bend-induced loss, 31, 33
  calculation, 31
  factors affecting, 31, 277

Gigabit Ethernet link budget model, 299
glass types, 30, 65
Automatic mask testing, 348–349
Avalanche photodiode (APD)
characteristics, 149
overview, 147–148
structure, 148–149
superlattice devices, 150

Ball grid array (BGA), 234, 256
Bandgap energy, 92
Bandwidth, photodiode, 135, 137
Bandwidth-length product (BWLP), 516
BC, *see* Beam centrality
BD, *see* Bore inner diameter
Beam centrality (BC), 198–200, 203
Bend-induced loss, 31, 33
BER, *see* Bit error rate
BGA, *see* Ball grid array
Bias voltage, photodiode, 136, 140, 143
BiCMOS, receiver logic and drive circuitry, 164–165
Bit error rate (BER)
calculation, 272
coupled power considerations, 195
error detector, 366–369
ESCON thresholding, 541
eyewidth, 275, 287
optical input power relationship, 274
overview, 160–161, 271
radiation effects, 291
testing, 339
Bore inner diameter (BD), coupling performance effects, 198, 200
BPON, *see* Broadband passive optical network
Broadband passive optical network (BPON), 414
Broadcast-and-select network, architecture, 632–633
Bump processes, flip-chip interconnection, 226–227
Bump structures, tape automated bonding, 223
Burn-in, 262
BWLP, *see* Bandwidth-length product

Cable television, attenuated fiber cables, 73–74
Calibration, detector, 137–138
California, environmental regulation, 458–460
CD, *see* Compact disk; Connector body dimension
Ceramic packages, 231–232
Chappe optical telegraph, 4–5
Chemical vapor deposition (CVD), fiber fabrication, 64
China, environmental regulation, 458, 462, 465–467
Chip-on-board (COB), 257
Chirped power penalty, 294
Chromatic bandwidth, 279–280
Chromatic dispersion, 279, 281, 298, 390–391
Classical IP over ATM, implementation, 492–493
Clock regeneration, fiber-optic transceiver, 243
CMOS, *see* Complementary metal oxide semiconductor
Coarse wavelength-division multiplexing (CWDM), 371
Coatings, optical fiber, 37
COB, *see* Chip-on-board
Code division multiple access (CDMA), passive optical network, 410
Color coding, TIA/EIA-598-A standard, 329
Compact disk (CD), lasers, 103–104, 118
Complementary metal oxide semiconductor (CMOS)
ESCON link design, 542
receiver logic and drive circuitry, 164–165
silicon photonics, *see* Silicon photonics
Computer clusters
copper versus optical technology tradeoffs, 429–431

geographically dispersed parallel sysplex features, 433–441
large-scale computing, 432–433
optical interconnects
case study, 453–454
Deep Computing, 448
hierarchy, 446
parallel supercomputers, 447–448
prospects, 448–450
parallel sysplex architecture, 433–441
supercomputer speeds, 433
switched interconnect fabric, 430–431
symmetric multiprocessor network, 427–428
system area network characteristics, 428–429
time synchronization, 442–446
topologies, 431–432
Connector body dimension (CD), 203
Connectorized module, 194
Connectors, *see also* Optical backplane
classification, 37–38
curing, 336–337
ESCON connectors
multimode, 549, 551
single mode, 553–555
FC connectors, 38
field and factory connectorization, 335–337
installation loss, 277
losses, 39
preconnectorized assemblies, 332–333
SC connectors, 38–39, 43
SFF interfaces, *see* Small form factor fiber-optic interfaces
specialty fibers, 66–67
Coordinated timing network (CTN), 444–446
Coordinated Universal Time (UTC), 442
Coupled power range (CPR)
overview, 194–195
variables affecting, 203
Coupling efficiency
equations, 27–29
examples, 28–29
Coupling loss, equations, 27, 30
CPR, *see* Coupled power range
CRC, *see* Cyclic redundancy check
Cross-plug range, 197
CTN, *see* Coordinated timing network
Current, photodiode equation, 134–135
Cutoff wavelength, 24
CVD, *see* Chemical vapor deposition
CWDM, *see* Coarse wavelength-division multiplexing
Cyclic redundancy check (CRC), 573

Dark current, detector, 133–134, 155
Data center, link planning and building, 320–321
Data pattern, construction for transceiver testing, 340–341
DBR, *see* Distributed Bragg reflector
DC, *see* Directional coupler
DDM, *see* Dense wavelength-division multiplexing
Deep Computing, 448
Demultiplexer, fiber-optic transceiver, 243
Dense wavelength-division multiplexing (DWDM), 80–81, 371, 437
Density of the states (DOS), nanophotonics, 720
Deserializer, fiber-optic transceiver, 243
Detectivity, equations, 137
Detector
avalanche photodiode, 147–150
metal-semiconductor-metal detector, 153–154
noise
amplification, 157–158
overview, 155–157
shot noise, 158
signal-to-noise ratio, 160–161
sources, 159–160
thermal noise, 158–159
optical subassembly coupling from fiber, 183–185
photodiode array, 150
PIN photodiode, 138–146

resonant-cavity enhanced photodetector, 154–155
Schottky barrier photodiode, 151–153
terminology and characteristics, 133–138
Device under test (DUT), 251
Die attach, wire bond, 222
Differential mode delay (DMD), 71–72, 180
Differential quantum efficiency, 98
Diffusion length, equation, 95
Digital video disk (DVD), lasers, 108–110
DIP, *see* Dual in-line package
Direct attach storage device (DASD), 433, 439, 441
Directional coupler (DC), silicon photonics, 687–688
Dispersion
equations, 32, 34
factors affecting, 31–32
link budget analysis, 279–281
modal dispersion, 34, 36
multimode fiber, 36
Dispersion-flattened fiber, 79–81
Dispersion-optimized fiber, 79
Dispersion penalty, 279–281
Dispersion-shifted fiber (DSF), 78–79
Distributed Bragg reflector (DBR)
mirrors, 114–115
tunable transmitter mechanisms, 636
vertical cavity surface-emitting laser, 111–115
Distributed Bragg reflector, 111–115
DMD, *see* Differential mode delay
DO, *see* Density of the states
Doped-fiber amplifier, 642
Doping, concentrations, 92
Driver
circuitry, 167–168
electrical interface, 165–166
optical interface, 166
system overview, 163–165
DSD, *see* Direct attach storage device
DSF, *see* Dispersion-shifted fiber
Dual in-line package (DIP), 230
Duct, utilization in cable installation, 332
DUT, *see* Device under test
DVD, *see* Digital video disk
Dynamic range, detector, 137

EDC, *see* Electronic dispersion compensation
EDFA, *see* Erbium doped fiber amplifier
Edge emitting light-emitting diode (E-LED)
heterojunction diode, 96
output, 96
principles, 91–99
Edge-emitting laser, 102–110, 182, 265–266
EFA, *see* Fiber end-face angle
EFMA, *see* Ethernet First Mile Alliance
Electric field, equations, 19
Electromagnetic compatibility (EMC), fiber-optic transceiver, 250, 252–253
Electromagnetic immunity (EMI), fiber-optic transceiver, 250, 252–253
Electronic dispersion compensation (EDC), 519
Electrostatic discharge (ESD), immunity in fiber-optic transceivers, 253–254
E-LED, *see* Edge emitting light-emitting diode
EMBH laser, *see* Etched-mesa buried-heterostructure laser
EMC, *see* Electromagnetic compatibility
EMI, *see* Electromagnetic immunity
Enhanced Ethernet, 587–588
Enterprise Systems Connection (ESCON)
architecture, 537–541
bit error rate thresholding, 541
data encoding/decoding, 540–541
geographically dispersed parallel sysplex, 438–439, 441
historical perspective, 537
link budget calculation case study, 567–569

link design
multimode
design and specification, 544–545
ESCON connector, 549, 551
fiber optic cable specifications, 548–550
input optical interface, 547
major components, 542–543
output optical interface, 546–547
single mode
design and specification, 546
overview, 543–544
link protocol, 538–540
loss budget analysis
multimode cable plant link loss, 560–562
single-mode cable plant link loss, 562–563
planning and installation, 555–560
prospects, 563–565
SBCON standard, 563
single-mode physical layer
fiber optic cable specifications, 551–553
input optical interface, 551
output optical interface, 551
XDF connector, 553–555
topology, 538
Environmental regulation
California, 458–460
China, 458, 462, 465–467
compliance, 469–470
European Union, 455–458, 463, 465–466, 468
industry impact, 470–471
Japan, 460–462
overview, 455–457
Restriction of Hazardous Substances, 463–469
South Korea, 462–463
EOM, *see* External optical modulator
EPON, *see* Ethernet passive optical network
ER, *see* Extinction ratio
Erbium doped fiber amplifier (EDFA)
advantages, 293–294
birefringence, 77
silicon photonics, 703–704
wavelength-division multiplexing, 383–385
Error detector, transceiver testing, 366–369
ESCON over Fibre Channel, *see* FICON
ESCON, *see* Enterprise Systems Connection
ESD, *see* Electrostatic discharge
Etched-mesa buried-heterostructure (EMBH) laser, 108
Ethernet
auto-negotiation, 575
carrier attributes, 745
Enhanced Ethernet, 587–588
fiber-optic attachment options, 745–749
40 Gigabit Ethernet, 582
Gigabit Ethernet, 581, 741
historical perspective, 571
metropolitan area networks, 583–586, 744
100 Gigabit Ethernet, 582
overview, 572–575
physical layer
copper links, 576–578
Fast Ethernet, 579, 741
100BASE-FX, 580–581
100BASE-T4, 579–580
100BASE-TX, 580
100BASE-X, 580
standards, 577, 740–742
roadmaps, 583
10 Gigabit Ethernet, 582, 742–743
triple play network case study, 591–592
Ethernet First Mile Alliance (EFMA), 586, 743–744
Ethernet passive optical network (EPON)
overview, 417–419, 586–587
10G Ethernet passive optical network, 419–420
ETR, *see* External timing reference
EU, *see* European Union

European Union (EU), environmental regulation, 455–458, 463, 465–466, 468
Evanescent fields, nanophotonics, 715–716
Extended remote copy (XRC), 440, 445
External differential quantum efficiency, 98
External optical modulator (EOM), 372
External timing reference (ETR), geographically dispersed parallel sysplex, 435, 437–438
Extinction ratio (ER)
  laser transmitter, 355–356
  measurement accuracy, 356–358)
  overview, 284–285
Eye mask
  dimensions and coordinate systems, 346–348
  failure cause diagnosis, 349–350
  shapes, 350–351
  testing, 346
Eye-diagram, construction for transceiver testing, 341–344
Eyewidth, 275, 287

Fabrication
  nanofiber, 721–722
  optical fiber, 63–64
Fabry-Perot (FP) cavity
  resonant-cavity enhanced photodetector, 154–155
  tunable receiver mechanisms, 637–638
Fabry-Perot amplifier (FPA), 641
Failures in time (FIT), 259
Fast Ethernet, 579, 741
Fast link pulse (FLP), 575
FBG, *see* Fiber Bragg grating
FBS, *see* Fiber beam spot
FCA, *see* Free carrier absorption
FCE, *see* Ferrule/core eccentricity
FCS, *see* Frame check sequence
FD, *see* Ferrule diameter
FDDI, *see* Fiber distributed data interface
FEC, *see* Forward error correction
Ferrule/core eccentricity (FCE), 199–200, 203
Ferrule diameter (FD), coupling performance effects, 198, 200, 203
Ferrule float (FF), 202–203
FF, *see* Ferrule float
Fiber beam spot (FBS), 197, 203
Fiber Bragg grating (FBG), 82, 374
Fiber distributed data interface (FDDI)
  FDDI-II, 603
  historical perspective, 13–14
  MAC specifications
    beacon process, 600
    claim token process, 600
    frame, 597–598
    ring operation, 598
    ring scheduling, 599–600
    token, 596–597
  physical layer
    overview, 593–594
    specifications
      coding and symbol set, 595
      connector types, 595
      interface power levels, 595
      line states, 596
  standards, 735–736
  station management
    physical connection management, 602–603
    SMT header, 601
    SMT information, 601–602
  station types, 593–594
Fiber end-face angle (EFA), 201–203, 206
Fiber image guide (FIG), 670
Fiber-Jack (FJ) connector, 55–56
Fiber optics
  cable handling
    duct utilization, 332
    maximum tensile load, 330–331
    maximum vertical rise, 331–332
    minimum bend radius, 330–331
    preconnectorized assemblies, 332–333
    protection of cable, 332

slack, 333
splicing methods, 333–334
characterization of fibers
attenuation, 30–31
dispersion, 31–32, 34, 36
materials, 30
mechanical properties, 36
coatings, 37
fabrication, 63–64
geometry of fiber, 20
historical perspective
data communications, 12–14
digital communications, 11–12
Internet, 14–17
optical systems testing and building, 10–11
origins, 7–10
Fiber-optic transceiver (TRX)
multisource agreements, 241–242
noise testing
device under test description, 251
electromagnetic compatibility
emission, 252
immunity, 252–254
noise on Vcc, 251–252
performance requirements, 250–251
optical interface
connectors and active device receptacles, 246, 248
optical fiber, 248–250
packaging
basic considerations, 254–255
motherboard assembly techniques, 255–256
open module design, 256–257
small form factor pluggable design, 258
washable sealed module design, 256
parallel fiber-optic transceiver
advantages, 245
transponders, 245–246
parallel optical links
high-density point-to-point communications, 266–267
link reach, 268–269
optic module configurations, 267–268
prospects, 269–270
physical layer interface, 242–243
prospects
demand, 262–263
functional integration, 264
geometrical outline, 263
mode underfill, 265–266
optical sources, 264–266
overview, 752–753
serial fiber-optic transceiver, 244–245
series production
basic considerations, 258–260
burn-in, 262
correlated environmental tests on subcomponents, 259
frozen process, 260
machine capability, 260
process capability, 259
statistical process control and random sampling, 260
testing
automatic mask testing, 348–349
data pattern construction, 340–341
error detector, 366–369
extinction ratio
overview, 355–356
measurement accuracy, 356–358
eye mask
dimensions and coordinate systems, 346–348
failure cause diagnosis, 349–350
shapes, 350–351
testing, 346
eye-diagram construction, 341–344
Golden PLL testing, 360–361
jitter analysis, 361, 363–364
oscilloscope frequency response, 345–346
pattern generator, 365–366
receiver test, 364
reference receiver, 351–354
Fiber Quick Connect (FQC), 314–315
Fiber Transport Services (FTS), 314–316

Fibre Channel
  historical perspective, 505
  overview, 506–510
  layers, 507–508
  network channels, 508–509
  architecture, 510
  data rates, 510–511
  data structures, 511–512
  roadmap branches, 513–515
  multimode link considerations, 515–519
  nomenclature, 516
  link power budget estimation, 519–522
  single-mode link considerations, 522–523
  mapping to upper level protocols
    ESCON over Fibre Channel, *see* FICON
    FC-4, 524
    IP over Fibre Channel, 524–525
    SCSI over Fibre Channel, 525
  service classes, 526–527
  metropolitan area networks, 527–529
  wide area networks, 527–529
  prospects, 529–530
FICON
  mainframe redesign case study, 533–534
  overview, 70, 525–526
  storage area network extension for disaster recovery, 535–536
Field programmable gate array (FPGA), 670
FIG, *see* Fiber image guide
Figure of merit (FOM), silicon photonics, 679
Finesse, ring resonator, 696
FiOS, 422
FIT, *see* Failures in time
FJ connector, *see* Fiber-Jack connector
Flame-retardant polyethylene (FRPE), 314
Flip-chip interconnection
  attachment, 227–228
  bump processes, 226–227
  overview, 225–226
  reliability, 228
  substrates, 227
FLP, *see* Fast link pulse
FOM, *see* Figure of merit
Forward error correction (FEC), optical transport network, 387
40 Gigabit Ethernet, 582
Four-wave mixing (FWM), 78–79
FP cavity, *see* Fabry-Perot cavity
FPA, *see* Fabry-Perot amplifier
FPGA, *see* Field programmable gate array
FQC, *see* Fiber Quick Connect
Frame check sequence (FCS), 573
Free carrier absorption (FCA), silicon photonics, 704
Free-space coupling, 659–660, 673, 751
Frozen process, 260
FRPE, *see* Flame-retardant polyethylene
FTS, *see* Fiber Transport Services
FWM, *see* Four-wave mixing

G.709 standard, 385–389
Gain coefficient, laser, 98
Gain equalization, amplifier, 643
GBIC, *see* Gigabit Interface Card
GDPS, *see* Geographically dispersed parallel sysplex
Generalized multi-protocol label switching (GMPLS), 377, 653
Generic Frame Procedure (GFP)
  frame structure, 499–500
  modes of operation, 499
  standards, 498, 501
Geographically dispersed parallel sysplex (GDPS), features, 433–441
GFP, *see* Generic Frame Procedure
GGP fiber, *see* Glass-glass-polymer fiber
Gigabit Ethernet, 581, 741
Gigabit Ethernet link budget model, 295–299
Gigabit Interface Card (GBIC), 258
Gigabit PON (GPON), 415–717
Gigalink card (GLC), 257–258
Glass-glass-polymer (GGP) fiber, VF-45 connector, 50
Glass-reinforced plastic (GRP), 318
GLC, *see* Gigalink card

GMPLS, *see* Generalized multi-protocol label switching
Golden PLL testing, 360–361
GPON, *see* Gigabit PON
GPR, *see* Group protection ring
Graded-index multimode fiber, 265–266
Grid computing, wavelength-division multiplexing case study, 403–404
GRIN SQW laser, 105
Group protection ring (GPR), 381–382
Group velocity dispersion (GVD), 77
GRP, *see* Glass-reinforced plastic
GVD, *see* Group velocity dispersion

Handling, cable
  duct utilization, 332
  maximum tensile load, 330–331
  maximum vertical rise, 331–332
  minimum bend radius, 330–331
  preconnectorized assemblies, 332–333
  protection of cable, 332
  slack, 333
  splicing methods, 333–334
HC, *see* Horizontal cross-connect
Heterojunction diode, edge emitting light-emitting diode, 96
High-performance parallel interface (HIPPI), 736
HiPerLink, geographically dispersed parallel sysplex, 434, 437–438
HIPPI, *see* High-performance parallel interface
Horizontal cross-connect (HC), 308, 317
Host bus adapter (HBA), 267
Hybrid silicon laser, 706–707

IB, *see* InfiniBand
IC, *see* Intermediate cross-connect
InfiniBand (IB)
  fiber-optic cable plant specifications
    connectors and splices, 625–626
    optical fiber, 625–626
  jitter
    methodology, 618
    specification, 616–617
  link layer
    electrical interface, 608–609
    fiber plant technology solutions, 609–612
    lanes, 606
    packet format, 607–608
  optical link systems
    1X system, 619
    4X system, 619–620
    8X-SX system, 620
    12X-SX system, 620–621
  optical modulation amplitude, 616
  optical receptor and connector
    1X connector, 622
    4X-SX connector, 622–624
    8X-SX connector, 624
    12X-SX connector, 624–625
  optical signal
    bit-to-bit skew, 618
    polarity and quiescent condition, 613
    rise/fall time measurement, 614–615
    transceiver, 610
    transmitter mask compliance, 613–614
  overview, 605–606
  performance, 606
Infrared reflow (IR), 234–235
Installation, *see* Link planning and building
Interbuilding network
  applications, 317
  backbone, 316
  loose-tube cables, 318–320
  topologies, 316–317
Intermediate cross-connect (IC), 308, 317, 334–335
Internet, fiber optic booms, 14–16
Intersymbol interference (ISI), 295, 297–298
Intrabuilding network
  applications, 310
  backbone, 307–308
  Fiber Transport Services, 314–316
  horizontal cabling, 309–310

tight-buffered cables, 310, 312, 314–316
topologies, 308
IR, *see* Infrared reflow
ISI, *see* Intersymbol interference

Jacket
specification, 66
tight-buffered cables, 316
Japan, environmental regulation, 460–462
Jitter
InfiniBand
methodology, 618
specification, 616–617
link budget analysis, 287–289
sources, 288
transceiver testing, 361, 363–364
Jitter transfer function (JTF), 289
JTF, *see* Jitter transfer function

Kelvin, Celsius conversion, 729

LAD system, *see* Linear add/drop system
LAN, *see* Local area network
LANE, *see* Local area network emulation
Large effective area fiber (LEAF), 80
LARNET, *see* Local access router net
Laser, *see also* Vertical cavity surface-emitting laser
edge-emitting laser, 102–110
Fibre Channel single-mode link, 522–523
optical communications history, 6–7
output, 96
principles, 91–99
safety
certification, 759–761
standards, 755–756, 758
silicon photonics
hybrid silicon laser, 706–707
linear cavity laser, 705
ring cavity laser, 705–706
structure, 100–102
threshold condition, 97
tunable transmitter mechanisms, 635–637
LC connector, 51–54
LCAS, *see* Link Capacity Adjustment Scheme
LCT, *see* Link confidence test; Liquid crystal technology
Leadframes, tape automated bonding, 223
LEAF, *see* Large effective area fiber
LED, *see* Light-emitting diode
Lens
magnification factor, 30
optical interconnect for packaging assembly, 237
Light propagation
multimode fiber, 26–27
single-mode fiber, 20–26
Light-emitting diode (LED)
blue devices, 110
characteristic curves, 99
optical subassembly coupling into fiber, 179–183
output, 96
principles, 91–99
structure, 95, 99
Limiting receiver, 173–175
Line overhead (LOH), SONET, 738
Line overhead octets, SONET, 479–480
Linear add/drop (LAD) system, 376–377
Linear cavity laser, silicon photonics, 705
Linearity, detector, 137
Linear receiver, 172–173, 175
Link budget analysis
ESCON calculation case study, 567–569
Fibre Channel link power budget estimation, 519–522
Gigabit Ethernet link budget model, 295–299
installation loss
attenuation versus wavelength, 277
connector loss, 277
splice loss, 277–278
transmission loss, 276–277

optical amplification link budgets, 299–300
optical power penalties
dispersion, 279–281
extinction ratio, 284–285
jitter, 287–289
modal noise, 290
mode partition noise, 282–284
multipath interference, 285–286
nonlinear noise effects, 291–295
overview, 278–279
radiation-induced loss, 290–291
relative intensity noise, 286–287
wavelength-division multiplexing system, 305
Link Capacity Adjustment Scheme (LCAS), 501–502
Link confidence test (LCT), 603
Link planning and building
cable handling
duct utilization, 332
maximum tensile load, 330–331
maximum vertical rise, 331–332
minimum bend radius, 330–331
preconnectorized assemblies, 332–333
protection of cable, 332
slack, 333
splicing methods, 333–334
connectors, *see* Connectors
data center, 320–321
indoor hardware
intermediate cross-connect, 334–335
telecommunications closet, 335
work area telecommunications outlet, 335
private networks, *see* Interbuilding network; Intrabuilding network
standards
National Electrical Code, 322–323
TIA/EIA-598-A, 329
TINEIA-568-A, 323–329
Liquid crystal technology (LCT), 383
LLC, *see* Logical link control
Local access router net (LARNET), 421
Local area network (LAN)
fiber distributed data interface, *see* Fiber distributed data interface
standards, 735–736
Local area network emulation (LANE)
components, 494
definition, 493
LAN emulation client
control connections, 495
data connections, 495–496
initialization phases, 494
operations, 496–498
objectives, 493
Lock-in amplifier, 157
Logical bus, 325
Logical link control (LLC), 573–574
LOH, *see* Line overhead
Loose-tube cables, 318–320
Loss budget analysis, ESCON
multimode cable plant link loss, 560–562
single-mode cable plant link loss, 562–563
LP mode, 22–24

MAC, *see* Media Access Control
Mach-Zehnder filter, tunable receiver mechanisms, 638–639
Mach-Zehnder interferometer (MZI), silicon photonics, 690–693
Machine capability, 260
Main cross-connect (MC), 308, 316–317
Main distribution facility (MDF), 315
Maser, historical perspective, 6
Massively parallel processor (MPP), 432
Materials, optical fiber, 30
Maximum tensile load, cable installation, 330–331
Maximum tolerable input jitter (MTIJ), 289
Maximum transmission unit (MTU), 573–574
Maximum vertical rise, cable installation, 331–332
MBE, *see* Molecular beam epitaxy

MC, *see* Main cross-connect
MCP, *see* Mode conditioning patch
MCVD, *see* Modified chemical vapor deposition
MDF, *see* Main distribution facility
MDX, *see* Multiplexing/demultiplexing
Mechanical properties, optical fiber, 36
Media Access Control (MAC)
  Ethernet, 572–573
  fiber distributed data interface specifications
    beacon process, 600
    claim token process, 600
    frame, 597–598
    ring operation, 598
    ring scheduling, 599–600
    token, 596–597
MEF, *see* Metro Ethernet Forum
MER, *see* Modulation error ratio
Metal packages, 229
Metal-semiconductor-metal (MSM) detector, 151, 153–154
Metric system
  conversion tables, 729–730
  prefixes, 730
Metro Ethernet Forum (MEF), 584, 744
Metropolitan area network
  Ethernet, 583–586, 744
  Fibre Channel over networks, 527–529
MFD, *see* Mode field diameter
Microresonator, nanofiber, 725–726
Minimum bend radius, cable installation, 330–331
MMI coupler, *see* Multimode interference coupler
Modal bandwidth, 279
Modal noise, optical power penalty, 290
Mode conditioning patch (MCP), cables, 70, 72–73
Mode field diameter (MFD), equation, 24
Mode partition noise, 273, 282–284
Mode underfill, 265–266
Modified chemical vapor deposition (MCVD), fiber fabrication, 64
Modulation error ratio (MER)
  advantages over bit error rate, 275–276
  calculation, 275
  definition, 275
Module alignment, 202
Molecular beam epitaxy (MBE), 102
MPP, *see* Massively parallel processor
MQW, *see* Multiquantum well
MSA, *see* Multisource agreement
MSM detector, *see* Metal-semiconductor-metal detector
MTIJ, *see* Maximum tolerable input jitter
MTP connector, *see* Multifiber termination push-on connector
MT-RJ connector, 44–47
MTU, *see* Maximum transmission unit
MU connector, 56–57
Multifiber termination push-on (MTP) connector, 315
Multimode fiber
  connections versus single-mode fiber, 196
  dispersion, 36
  light propagation, 26–27
  number of propagating modes, 26
  reuse for high-speed storage area networks, 87–90
Multimode interference (MMI) coupler, silicon photonics, 688–689
Multipath interference, 285–286
Multiplexer, fiber-optic transceiver, 243
Multiplexing/demultiplexing (MDX), filters, 372
Multiplication factor, avalanche photodiode, 149
Multiquantum well (MQW), laser, 104
Multisource agreement (MSA)
  fiber-optic transceiver, 241–242
  small form factor fiber-optic interfaces
    QSFP, 59–60
    SFP, 57–58
    SFP Plus, 58
Muxponder, 372
Myrinet, 432
MZI, *see* Mach-Zehnder interferometer

NA, *see* Numerical aperture
Nanofiber
  device applications, 724–726
  fabrication, 721–722
  waveguiding properties, 722–724
Nanophotonics
  definition, 713
  evanescent fields, 715–716
  historical perspective, 713–714
  near-field, 714–715
  quantum confined effects, 719–720
  surface plasmon resonance, 717–719
National Electrical Code (NEC), 322–323
National LambdaRail Project, 399–401
Near end crosstalk (NEXT), wavelength-division multiplexing, 78
Near-field, nanophotonics, 714–715
NEC, *see* National Electrical Code
NEP, *see* Noise equivalent power
NEXT, *see* Near end crosstalk
Noise
  detector
    amplification, 157–158
    overview, 155–157
    shot noise, 158
    signal-to-noise ratio, 160–161
    sources, 159–160
    thermal noise, 158–159
  fiber-optic transceiver testing
    device under test description, 251
    electromagnetic compatibility
      emission, 252
      immunity, 252–254
    noise on Vcc, 251–252
    performance requirements, 250–251
  nonlinear noise effects on link budget, 291–295
Noise equivalent power (NEP), detector, 136
Noise floor, detector, 135
Nonuniform memory architecture (NUMA), 453
Nonzero-dispersion-shifted fiber (NZDSF), 78
NUMA, *see* Nonuniform memory architecture
Numerical aperture (NA), 27
NZDSF, *see* Nonzero-dispersion-shifted fiber

O/E converter, *see* Optical-to-electrical converter
OADM, *see* Optical add/drop multiplexer
OAN, *see* Optical access network
O-BLSR, 379
OC-48, 631
OC-768, 632
OCF-192, 631–632
OCh-DPRing, 379–382
OCh-SPRing, 379–381
ODN, *see* Optical distribution network
OFB, *see* Ordered fiber bundle
OFC, *see* Open fiber control
OFNP, *see* Optical fiber nonconductive plenum
OFNR, *see* Optical fiber nonconductive riser
OLT, *see* Optical line termination
OM1 fiber, *see* Optical Multimode 1 fiber
OM2 fiber, *see* Optical Multimode 2 fiber
OM3 fiber, *see* Optical Multimode 3 fiber
OMA, *see* Optical modulation amplitude
100BASE-FX Ethernet, 580–581
100BASE-T4 Ethernet, 579–580
100BASE-TX Ethernet, 580
100BASE-X Ethernet, 580
100 Gigabit Ethernet, 582
ONU, *see* Optical network unit
Open fiber control (OFC), 434
Open Systems Interconnect (OSI), 572, 733
Optical access network (OAN), passive optical networks, 407
Optical add/drop multiplexer (OADM), 373, 696
Optical alignment, 197
Optical amplifiers
  doped-fiber amplifier, 642
  gain equalization, 643

semiconductor optical amplifier, 641–642
silicon photonic Raman amplifiers and lasers, 704–707
Optical backplane
free-space coupling, 659–660, 673
integrated polymer waveguides, 661–663
overview, 658–659
printed circuit board coupling, 660
prospects, 752
Optical board interconnects
examples, 664–666
intraboard waveguide coupling, 663–664
Optical channel payload unit (OPU), 386
Optical chip interconnection
examples, 667–671
intrachip communication, 671–672
overview, 666–667
Optical distribution network (ODN), 408–409
Optical fiber nonconductive plenum (OFNP), 310, 322
Optical fiber nonconductive riser (OFNR), 310, 322
Optical fiber, *see* Fiber optics
Optical interconnect
computer clusters
case study, 453–454
Deep Computing, 448
hierarchy, 446
parallel supercomputers, 447–448
prospects, 448–450
fiber-optic transceiver
connectors and active device receptacles, 246, 248
optical fiber, 248–250
packaging assembly
coupling, 235–236
lenses, 237
prisms, 237
waveguides, 236–237
Optical line termination (OLT), 408–409
Optical modulation amplitude (OMA)
InfiniBand, 616
overview, 285
transceiver testing, 359
Optical Multimode 1 (OM1) fiber, 67
Optical Multimode 2 (OM2) fiber, 67–68
Optical Multimode 3 (OM3) fiber, 69, 88
Optical network unit (ONU), 408
Optical power, equations, 95, 97
Optical signal-to-noise ratio (OSNR), 299–300
Optical splitter, 408
Optical subssembly (OSA)
fiber radiation coupling into photodetector, 183–185
function, 177
laser diode radiation coupling into fiber, 179–183
packaging, 185–188
parallel optical links, 188–190
prospects, 190–191
receiver optical subssembly properties, 179
transmitter optical subssembly properties, 178–179
Optical telegraph, historical perspective, 3–6
Optical transport channel unit (OTU), 386
Optical transport network (OTN)
applications, 388–389
forward error correction, 387
G.709 standard, 385
layers, 386
monitoring, 387–388
SONET/SDH internetworking, 388
Optical transport section (OTS), 386
Optical User Network Interface (O-UNI), 653
Optical-to-electrical (O/E) converter, 355
OPU, *see* Optical channel payload unit
Ordered fiber bundle (OFB), 670
OSA, *see* Optical subssembly
Oscilloscope, transceiver testing
automatic mask testing, 348–349
data pattern construction, 340–341

extinction ratio
measurement accuracy, 356–358
overview, 355–356
eye mask
dimensions and coordinate systems, 346–348
failure cause diagnosis, 349–350
shapes, 350–351
testing, 346
eye-diagram construction, 341–344
Golden PLL testing, 360–361
jitter analysis, 361, 363–364
oscilloscope frequency response, 345–346
reference receiver, 351–354
OSI, *see* Open Systems Interconnect
OSNR, *see* Optical signal-to-noise ratio
OTN, *see* Optical transport network
OTS, *see* Optical transport section
OTU, *see* Optical transport channel unit
O-UNI, *see* Optical User Network Interface
Outside vapor deposition (OVD), fiber fabrication, 64
OVD, *see* Outside vapor deposition

Packaging assembly
board attachment
pinned package, 232–233
surface mount package, 234–235
first-level interconnects
flip-chip interconnection
attachment, 227–228
bump processes, 226–227
overview, 225–226
reliability, 228
substrates, 227
tape automated bonding
bonding, 224
bump structures, 223
leadframes, 223
overview, 222–223
reliability, 224
wire bond
die attach, 222
reliability, 222
thermocompression wire bond, 220–221
thermostatic wire bond, 220–221
ultrasonic wire bond, 221–222
optical interconnect
coupling, 235–236
lenses, 237
prisms, 237
waveguides, 236–237
requirements, 219–220
types
ceramic packages, 231–232
metal packages, 229
plastic packages, 229–231
PANDA fiber, *see* Polarization maintaining and absorption reducing fiber
Parallel fiber-optic transceiver
advantages, 245
transponders, 245–246
Parallel optical interconnects (POIs), detector arrays, 147
Parallel optical links
computer cluster optical interconnects
Deep Computing, 448
hierarchy, 446
parallel supercomputers, 447–448
prospects, 448–450
supercomputer case study, 453–454
high-density point-to-point communications, 266–267
link reach, 268–269
optic module configurations, 267–268
prospects, 269–270
Parallel optoelectrical interface
optical subssembly, 188–190
overview, 163–164
Parameter management frame (PMF), 602
Passive optical network (PON)
applications, 405
ATM passive optical network, 413–414
broadband passive optical network, 414
comparison of approaches, 423–424
Ethernet passive optical network, 417–419

FiOS, 422
gigabit PON, 415–717
optical access networks, 407
popularity, 406–407
principles, 405–406
standards and variants, 413
structure and function, 407–409
upstream access, 409–412
10G Ethernet passive optical network, 419–420
wavelength-division multiplexing passive optical network, 420–422
Path overhead octets, SONET, 479–480
Pattern generator, transceiver testing, 365–366
PBRS, *see* Pseudo-random binary sequence
PBX, *see* Private branch exchange
PCM, *see* Physical connection management
PCVD, *see* Plasma-assisted chemical vapor deposition
PDG, *see* Polarization-dependent gain
PDL, *see* Polarization-dependent loss
Peer-to-peer optical networking, 653–655
PGA, *see* Pin-grid array
Phase-locked loop (PLL), fiber-optic transceiver, 243
Photocurrent, avalanche photodiode, 149
Photodiode, *see* Detector
Photodiode array, 150
Photosensitive fiber, 81–82
Physical connection management (PCM), 602–603
Physical constants, 731
Physical star
advantages and disadvantages, 327–328
implementation, 328–329
PIN detector, vertical cavity surface-emitting laser, 122–123
PIN photodiode, *see* Positive-intrinsic-negative photodiode
Pin-grid array (PGA), 231
Pinned package, 232–233
Planar alignment, 198–200
Planck's Law, 94
Plasma-assisted chemical vapor deposition (PCVD), fiber fabrication, 64
Plastic lead chip carrier (PLCC), 234
Plastic packages, 229–231
PLCC, *see* Plastic lead chip carrier
PLL, *see* Phase-locked loop
Plug repeatability (PR), 197
PMF, *see* Parameter management frame; Polarization maintaining fiber
pn photodiode, features, 138
Point-to-point, logical topology, 325
POIs, *see* Parallel optical interconnects
Polarization maintaining and absorption reducing (PANDA) fiber, 76
Polarization maintaining fiber (PMF), 74–77
Polarization-dependent gain (PDG), 77
Polarization-dependent loss (PDL), 77
Polarized-mode dispersion fiber, wavelength-division multiplexing applications, 391
PON, *see* Passive optical network
Positive-intrinsic-negative (PIN) photodiode
circuitry, 145
principles, 138–139
quantum efficiency, 141–143
sample specifications, 143–144
self-capacitance, 140
semiconductor properties, 141
structure, 139–140
temperature effects, 143
Power coupling efficiency, 684–685
Power, accepted by fiber, 27
PR, *see* Plug repeatability
Prism, optical interconnect for packaging assembly, 237
Private branch exchange (PBX), 307–308, 654
Process capability, calculation, 259
Propagation constant, 22–23
Pseudo-random binary sequence (PBRS), 365–366

QE, *see* Quantum efficiency
QFP, *see* Quad flat pack
QSFP, 59–60
Quad flat pack (QFP), 234
Quantum confinement, silicon photonics, 703
Quantum efficiency (QE), detector, 134, 141–143

Radiation-induced loss, 290–291
RARP, *see* Reverse address resolution protocol
REACH directive, 464, 471
RECAP, *see* Resonant-cavity enhanced photodetector
Receiver, *see also* Fiber-optic transceiver
  circuitry, 168–172
  electrical interface, 165–166
  limiting receiver, 173–175
  linear receiver, 172–173, 175
  optical interface, 166
  optical subssembly properties, 179
  system overview, 163–165
  testing in transceiver, 364–365
  tunable receiver mechanisms, 637–641
Recombination, electrons and holes, 93–94
Reference receiver, 351–354
Refractive index
  calculation, 34
  profiles of fibers, 20–22, 79
Regenerator section overhead (RSOH), SONET, 738
Relative intensity noise (RIN), 273, 286–287, 298
Remote interrogation of terminal network (RITENET), 421–422
Resonant-cavity enhanced photodetector (RECAP), 154–155
Responsivity, detector, 133–134
Restriction of Hazardous Substances (RoHS), 463–469
Reverse address resolution protocol (RARP), 492
RIN, *see* Relative intensity noise
Ring, logical topology, 325
Ring cavity laser, silicon photonics, 705–706
Ring resonator (RR), silicon photonics, 695–697
RITENET, *see* Remote interrogation of terminal network
ROADM, 392–396
Rod and tube casting, fiber fabrication, 65
RoHS, *see* Restriction of Hazardous Substances
RR, *see* Ring resonator
RSOH, *see* Regenerator section overhead

SAM, *see* Subassembly alignment
SAMS, *see* Separation and multiplication layers
SAN, *see* Storage area network
Saturation, detector, 137
SBS, *see* Source beam spot; Stimulated Brillouin scattering
SC-DC connector, 47–49
Schottky barrier photodiode, 147, 151–153
SCMA, *see* Sub-carrier multiple access
SC-QC connector, 47–48
SCSI, *see* Small computer systems interface
SD, *see* Shroud dimension
SDM, *see* System data mover
Search for Extra-Terrestrial Intelligence (SETI), 447
Section overhead octets, SONET, 479
Self-capacitance, photodiode, 140
Sellmeier equation, 34
Semiconductor optical amplifier, 641–642
Separation and multiplication layers (SAMs), avalanche photodiode, 149–150
SERDES, *see* Serializer/deserializer
Serial fiber-optic transceiver, 244–245
Serializer, fiber-optic transceiver, 243
Serializer/deserializer (SERDES), 506
Serial optoelectrical interface, 163–164

Server time protocol (STP), 443–444
SETI, *see* Search for Extra-Terrestrial Intelligence
SFF fiber-optic interfaces, *see* Small form factor fiber-optic interfaces
SFP, 57–58, 258
SFP Plus, 58
Shielded twisted pair, standards, 578
Shot noise, detector, 158
Shroud dimension (SD), 203
Signal-to-noise ratio (SNR)
  bit error rate relationship, 272
  detector, 160–161
Silicon photonics
  active devices
    amplifiers and lasers
      overview, 702–704
      Raman amplifiers and lasers, 704–707
    modulators
      PN carrier depletion silicon, 40 Gb/s modulator, 699–701
      silicon metal-oxide-semiconductor modulator, 698–699
    near-infrared Si-Ge photodetector, 701–702
    refractive index control, 697–698
  advantages, 678
  applications, 680–681
  limitations, 678–679
  overview, 677–680
  passive devices
    couplers
      directional coupler, 687–688
      multimode interference coupler, 688–689
    Mach-Zehnder interferometer, 690–693
    ring resonator, 695–697
    waveguide Bragg grating, 693–695
    waveguide coupling, 684–686
    waveguides, 681–684
    Y-junction, 689–690
Single quantum well (SQW), laser, 104–105
Single-mode fiber
  connections versus multimode fiber, 196
  light propagation, 20–26
Slack, cable installation, 333
Small computer systems interface (SCSI), 506, 508
  over Fibre Channel, 525
Small form factor (SFF) fiber-optic interfaces
  comparison of connectors, 60–62
  Fiber-Jack connector, 55–56
  LC connector, 51–54
  MT-RJ connector, 44–47
  MU connector, 56–57
  multisource agreements
    QSFP, 59–60
    SFP, 57–58
    SFP Plus, 58
  SC-DC connector, 47–49
  SC-QC connector, 47–48
  standards, 43–44, 57
  VF-45 connector, 49–51
SMP network, *see* Symmetric multiprocessor network
SMT, *see* Surface mount technique
Snell's Law, 201
SNR, *see* Signal-to-noise ratio
SONET, *see* Synchronous optical network
Source beam spot (SBS), 197, 203
South Korea, environmental regulation, 462–463
SP, *see* Surface plasmon resonance
SPE, *see* Synchronous payload envelope
Splice, installation loss, 277–278
Splicing, methods, 333–334
Splitter, manufacture, 67
SQW, *see* Single quantum well
SRS, *see* Stimulated Raman scattering
Star
  logical topology, 325
  physical star
    advantages and disadvantages, 327–328
    implementation, 328–329

Stimulated Brillouin scattering (SBS), 292–293
Stimulated Raman scattering (SRS)
  overview, 293
  silicon photonics, 704
Storage area network (SAN)
  extension for disaster recovery, 535–536
  Fibre Channel, *see* Fibre Channel
  multimode fiber reuse for high-speed storage area networks, 87–90
STP, *see* Server time protocol
STS-1, 476
Subassembly alignment (SAM), 202–203
Sub-carrier multiple access (SCMA), passive optical network, 410–411
Supercomputer, *see* Computer clusters
Supperlattice avalanche photodiode, 150
Surface mount technique (SMT), 234–235, 256
Surface plasmon resonance (SPR), nanophotonics, 717–719
Surface wave filter (SWF), fiber-optic transceiver, 243
SWF, *see* Surface wave filter
Switchable grating, tunable receiver mechanisms, 639, 631
Switched interconnect fabric, 430–431
Symmetric multiprocessor (SMP) network, 427–428
Synchronous optical network (SONET)
  case study, 503–504
  framing, 477–478
  historical perspective, 475–476
  international interoperability, 482
  line overhead octets, 479–480
  multiplexing, 481
  overview, 475
  path overhead octets, 479–480
  payload capacity comparison versus ATM, 490
  payload envelope pointer, 480–481
  physical layer specifications, 482–483
  section overhead octets, 479
  SONET/SDH, 375, 378–379, 388, 737–739
  STS-1data rates, 476
  synchronous payload envelope, 478, 480
  virtual tributaries, 481–482
Synchronous payload envelope (SPE), 478, 480
System data mover (SDM), 445–446

TAB, *see* Tape automated bonding
Tape automated bonding (TAB)
  bonding, 224
  bump structures, 223
  leadframes, 223
  overview, 222–223
  reliability, 224
TDM, *see* Time-division multiplexing
TDMA, *see* Time domain multiple access
Telecommunications, optical communications history, 11–12
Telecommunications closet, 335
10 Gigabit Ethernet, 582, 742–743
10 Gigabit Ethernet passive optical network, 419–420
TFF, *see* Thin-film filter
Thermal noise, detector, 158–159
Thermocompression wire bond, 220–221
Thermostatic wire bond, 220–221
Thin-film filter (TFF), 374–375
Threshold condition, laser, 97
Threshold current, temperature effects, 98
TIA/EIA-598-A, color coding standard, 329
Tight-buffered cables, 310, 312, 314–316
Time-of-day (TOD) clock, 443
Time-division multiplexing (TDM), 372, 375–376
Time domain multiple access (TDMA), passive optical network, 409
Time synchronization, distributed computing, 442–446
Time-varying resistance, metal-semiconductor-metal detector, 154

TINEIA-568-A
  cabling topologies and applications, 323–324
  logical topology
    logical bus, 325
    physical star
      advantages and disadvantages, 327–328
      implementation, 328–329
    point-to-point, 325
    ring, 325
    star, 325
    token bus, 326
  network topologies, 324
TO, *see* Transistor outline
TOD clock, *see* Time-of-day clock
Token bus, logical topology, 326
TP, *see* Two-photon absorption
Transceiver, *see* Fiber-optic transceiver
Trans-impedance amplifier (TZA), optical interface, 166, 171–172
Transistor outline (TO)
  fiber radiation coupling to photodetector, 184
  metal packages, 229
  optical subssembly packaging, 185–187
  receiver optical subssembly, 179
  transmitter optical subssembly, 180
Transistor-transistor logic (TTL), receiver logic and drive circuitry, 165
Transmission loss, 276–277
Transmitter, *see also* Fiber-optic transceiver
  optical subssembly properties, 178–179
  tunable transmitter mechanisms, 635–637
Traveling wave amplifier (TWA), 641
TRX, *see* Fiber-optic transceiver
TTL, *see* Transistor-transistor logic
TWA, *see* Traveling wave amplifier
Twisted pair, *see* Shielded twisted pair; Unshielded twisted pair
Two-photon absorption (TPA), silicon photonics, 704
TZA, *see* Trans-impedance amplifier

Ultrasonic wire bond, 221–222
Unshielded twisted pair (UTP)
  fiber optic advantages, 309–310
  standards, 578
UTC, *see* Coordinated Universal Time
UTP, *see* Unshielded twisted pair

VAD, *see* Vapor axial deposition
Vapor axial deposition (VAD), fiber fabrication, 64
Vapor phase reflow (VPR), 234–235
VCSEL, *see* Vertical cavity surface-emitting laser
Vertical cavity surface-emitting laser (VCSEL)
  advantages, 101
  applications, 118–121
  distributed Bragg reflector, 111–115
  fiber-optic transceivers, 264–266
  Fibre Channel, 506–507
  intracavity metal contact, 115–116
  microcavity structure, 117
  modulation speed, 125
  optical subssembly coupling into fiber, 180–181
  passive automated alignment, 214–216
  PIN detector, 122–123
  power and efficiency, 101–102
  principles, 91–99
  strained InGaAs laser, 113–114
  structure, 111–113, 122
  super-low-threshold microcavity type, 123, 125
  temperature insensitivity, 117–118
VF-45 connector, 49–51
V grooves, alignment, 210–212
Virtual concatenation, 501
Virtual private network (VPN), 653
VPN, *see* Virtual private network
VPR, *see* Vapor phase reflow

Wave equation, 19–20, 74
Waveguide
  integrated polymer waveguides, 661–663

intraboard waveguide coupling, 663–664
nanofiber properties, 722–724
optical interconnect for packaging assembly, 236–237
silicon photonics, *see* Silicon photonics
Waveguide Bragg grating, silicon photonics, 693–695
Wavelength converter
optical gating wavelength conversion, 647–649
optoelectronic conversion, 647
wave-mixing wavelength conversion, 649
Wavelength domain multiple access (WDMA)
historical perspective, 633–635
passive optical network, 410
Wavelength-division multiplexing (WDM), *see also* Coarse wavelength-division multiplexing; Dense wavelength-division multiplexing
amplifiers
erbium-doped fiber amplifier, 383–385
optical transport network applications, 388–389
forward error correction, 387
G.709, 385
layers, 386
monitoring, 387–388
SONET/SDH internetworking, 388
attenuated fiber cables, 73–74
case studies
National LambdaRail Project, 399–401
optical networks for grid computing, 403–404
triple play network, 591–592
chromatic dispersion compensation, 390–391
detectors, 147, 150
fiber design, 80–81
filters, 373–375
40G applications, 389–390
historical perspective, 15–17, 633–635
link budget design case study, 305
metro and regional networks, 376–327
near end crosstalk, 78
optical switching, 382–383
polarized-mode dispersion fiber applications, 391
protection
group protection ring, 381–382
mechanisms, 377–379
O-BLSR, 379
OCh-DPRing, 379–381
OCh-SPRing, 379–381
ROADM, 392–396
system structure, 372
Wavelength-division multiplexing passive optical network (WDM PON), 420–422
Wavelength multiplexer/demultiplexer, 643–644
Wavelength-routed network
architecture, 633
router features, 644–646
Wavelength-selective switch (WSS), ROADM, 393–394
WDM PON, *see* Wavelength-division multiplexing passive optical network
WDM, *see* Wavelength-division multiplexing
WDMA, *see* Wavelength domain multiple access
WG, *see* Arrayed waveguide grating
Wide area network, Fibre Channel over networks, 527–529
Wigner-Kramer-Brillouin (WKB) approximation, 26–27
Wire bond
die attach, 222
reliability, 222
thermocompression wire bond, 220–221

thermostatic wire bond, 220–221
ultrasonic wire bond, 221–222
WKB, *see* Wigner-Kramer-Brillouin
Work area telecommunications outlet, 335
WSS, *see* Wavelength-selective switch

XDF connector, 553–555
XRC, *see* Extended remote copy

Y-junction, silicon photonics, 689–690